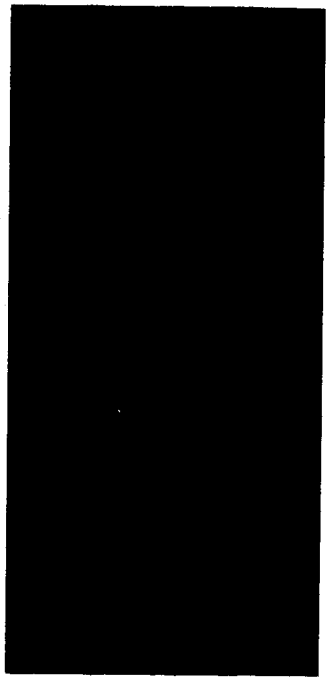
PERIODICA POLYTECHNICA
TRANSPORTATION ENGINEERING

TECHNICAL UNIVERSITY OF BUDAPEST



Vol. 24. No. 2.
1996

PERIODICA POLYTECHNICA

A contribution to international technical sciences, published by the Technical University of Budapest in six separate series, covering the following sciences:

Chemical Engineering
Civil Engineering
Electrical Engineering and Informatics
Mechanical Engineering
Social and Management Sciences (Earlier: Humanities and Social Sciences)
Transportation Engineering

TRANSPORTATION ENGINEERING SERIES

Published twice a year

Contributions in transportation engineering should be sent to the address of the journal:

TECHNICAL UNIVERSITY OF BUDAPEST
Periodica Polytechnica
Transportation Engineering
H-1521 Budapest, HUNGARY
Telefax: + 36 1 463-3041

For subscriptions please contact "Andreas Hess" Ltd. (Mail: P.O.B. 290, Budapest III, H–1300; Telefax: +36 1 250–2188) or its representatives abroad.

Exchange copies should be requested from the International Exchange Department of the Central Library of the Technical University of Budapest
(H–1111 Budapest, Budafoki út 4.; Telefax: + 36 1 463–2440).

Published by the Technical University of Budapest, Hungary, with the financial help of the foundation "Ipar a korszerű mérnökképzésért" (Industry for Modern Education in Engineering).

HU ISSN: 0303–7800

PERIODICA POLYTECHNICA

TRANSPORTATION ENGINEERING

Vol. 24 * No. 2 * 1996

TECHNICAL UNIVERSITY
BUDAPEST

PRINTED IN HUNGARY
LIGATURA LTD – ÁFÉSZ PRESS, VÁC

PERIODICA POLYTECHNICA SER. TRANSP. ENG. VOL. 24, NO. 2, PP. 81-94 (1996)

CLOSED LOOP IDENTIFICATION SCHEMES FOR ACTIVE SUSPENSION DESIGN

Péter GÁSPÁR, József BOKOR* and László PALKOVICS**

* Computer and Automation Research Institute
H-1111 Budapest, Kende u. 13-17
** Technical University of Budapest
H-1521 Budapest, Műegyetem rkp. 3

Received: Nov. 7, 1994

Abstract

This paper presents an overview of the most important closed loop identification methods, i.e. of the traditional direct/indirect, of the two stage and of the coprime factorization methods. The role of these algorithms in closed loop design is highlighted. These approaches are illustrated and compared with each other through a vehicle dynamical example taking the active suspension design problem.

Keywords: closed loop identification, identification for control, active suspension design.

1. Controller Design Based on Identified Model

The controller methods assume the knowledge of the actual plant. But in reality the transfer function of the plant is not known exactly, it is only known partially, and it has uncertainties and stochastic features. The model of the actual plant can be determined using measured input and output signals. The more exactly the model is known the more difficult it is to design and to implement the controller. So reduced complexity model is applied for the purpose of design a controller, which satisfies the stability and performance requirements and it can be implemented easily. The controller design based on identified model leads to iterative identification and controller design, where the model identification is performed in closed loop (GEVERS, 1991).

During the control design process the designer selects a controller so that it has to satisfy not only the stability but also some performance criteria. Let $P(q)$ be the transfer function of the actual plant, and let $C(q)$ mean the transfer function of the controller as it can be seen in *Fig. 1.*

The actual plant is modelled using the input and output signals, and the controller is designed on the basis of the identified model. The designed feedback loop consists of the identified model $\hat{P}(q,\theta)$ and the designed controller $C(q)$ as it can be seen in *Fig. 2.*

In reality the controller works together with the actual plant in the achieved closed loop system. So the controller has to be designed in a way that it has to be suitable not only for the identified model but for the actual

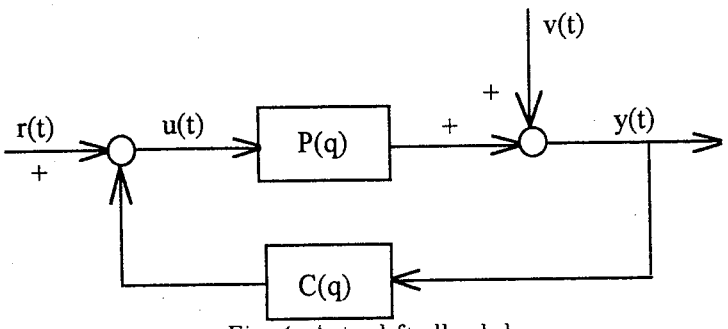


Fig. 1. Actual feedback loop

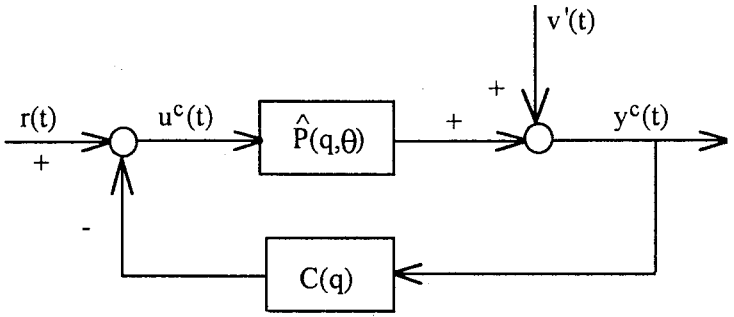


Fig. 2. Designed feedback loop

plant as well. This means that the controller has to ensure the stability of the achieved closed loop system in spite of the uncertainties of the actual plant, which is called as robust stability. On the other hand, the controller has to satisfy some prescribed performance criteria of the achieved closed loop system, which is called as robust performance.

This paper first highlights the theory of closed loop identification. Chapter 3 summarizes the traditional methods, while Chapter 4 describes the modern theories such as the two stage method and the coprime factorization method. Moreover Chapter 5 demonstrates an academic example for closed loop identification approaches.

2. Identification of Systems Operating in Closed Loop

If the input $u(t)$ is not a function of the output $y(t)$ then the plant is said to be operating in open loop, and the output does not feedback and does not affect the input. Unfortunately, practical system identification is usually performed while the plant is operating in a stable closed loop. Here the input is a function of the output and possible of an external loop input

$r(t)$. There are many reasons for identifying $\hat{P}(q,\theta)$ while it is operating in closed loop. If the plant is not stable in open loop operation, the closed loop control is required to maintain the stability for obvious reasons. If the plant is non-linear and the purpose of the identification is to develop a linear model about some nominal operating points, then the closed loop control has to maintain the plant near that nominal operating point.

The connections among the command $r(t)$, disturbance $v(t)$, input $u(t)$, and output $y(t)$ signals of the actual closed loop can be formalized as follows.

$$y(t) = [I + P(q)C(q)]^{-1}P(q)C(q)r(t) + [I + P(q)C(q)]^{-1}v(t)\,, \quad (1)$$

$$u(t) = C(q)[I + P(q)C(q)]^{-1}r(t) - C(q)[I + P(q)C(q)]^{-1}v(t)\,. \quad (2)$$

The identification cost function based on the predication error, $\varepsilon(t,\theta)$ which is as follows:

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t,\theta)\,, \quad (3)$$

where $\hat{y}(t,\theta) = \hat{P}(q,\theta)u(t)$. Then the prediction error cost function is as follows:

$$V_N(\theta) = \frac{1}{N}\sum_{t=1}^{N}[\varepsilon(t,\theta)]^2\,. \quad (4)$$

The formulas of the prediction error, where $\hat{y}(t,\theta)$ is expressed, in open loop case (5) and in closed loop case (6) are significantly different because of the feedback effect (LJUNG, 1987).

$$\varepsilon(t,\theta) = [P(q) - \hat{P}(q,\theta)]u(t) + v(t)\,, \quad (5)$$

$$\varepsilon(t,\theta) = [I + P(q)C(q)]^{-1}[P(q) - \hat{P}(q,\theta)]C(q)r(t) + [I - \hat{P}(q,\theta)C(q)][I + P(q)C(q)]^{-1}v(t)\,. \quad (6)$$

In the 70-s the direct and the indirect identification methods have been developed for closed loop identification. But the new control design approaches involve the development of modern closed loop identification methods, such as the two-stage method and the coprime factorization method.

3. Traditional Closed Loop Identification Methods

The direct identification neglects the effect of the feedback and the identification is performed using the input signal $u(t)$ and the output signal $y(t)$ in open loop way

$$y(t) = \hat{P}(q,\theta)u(t) + v(t)\,, \quad (7)$$

where $v(t)$ is an estimate of the noise signal. It can be proved that if the command signal $r(t)$ is persistently exciting of a sufficiently high order, then the estimate of $\hat{P}(q,\theta)$ is consistent. As a result a unique and consistent model is obtained despite of the presence of feedback. Moreover, it is sufficient that the external and persistently exciting signal is present (SÖDERSTRÖM, et al., 1976). The direct identification method generally results in a very complicated model.

As an alternative to the direct identification approach, the method of indirect identification avoids to take measured data from the closed loop. It assumes the knowledge of the controller and the measurability of the command signal $r(t)$. The indirect method consists of two steps. In the first step it computes the transfer function of the whole closed loop system $\hat{G}_c(q,\theta_c)$ using the command signal $r(t)$ and the output signal $y(t)$ in open loop way (see *Fig. 2*)

$$y(t) = \hat{G}_c(q,\theta_c)r(t) + w(t)\,, \tag{8}$$

where $w(t)$ is the noise signal. Then in the second step it determines the model applying matrix manipulations using the knowledge of the controller.

$$\hat{G}_c(q,\theta_c) = [1 + \hat{P}(q,\theta)C(q)]^{-1}\hat{P}(q,\theta)C(q)\,. \tag{9}$$

It can be proved that the conditions for a unique and consistent identification result are obtained as in the case of the direct identification with a persistently exciting external signal (SÖDERSTRÖM et al., 1976).

Conditions for consistency of the direct and of the indirect approaches are the same, but this does not mean that both methods give the same result in the finite sample case, or that they are equally easy to apply. Moreover, an advantage of the direct approach is that only one step is required. For the indirect approach it is not obvious how the equations in second step (9) should be solved.

4. Modern Closed Loop Identification Methods

4.1. The Two-Stage Method

The two-stage method is introduced in VAN DEN HOF and SCHRAMA, 1992. This method avoids complicatedly parametrized model sets, as are required in the direct method, and does not need apriori knowledge of the controller. It consists of two identification steps, which can be performed with open loop methods. If we investigate a single input, single output system, then let us define the $S(q)$ factor, as the common factor of the input and the output equations (1) and(2).

$$S(q) = C(q)[I + P(q)C(q)]^{-1}\,. \tag{10}$$

Since $r(t)$ and $v(t)$ are uncorrelated signals, moreover $u(t)$ and $r(t)$ are available from measurements, it follows that we can identify the $S(q)$ function in open loop way from the following equation:

$$u(t) = \hat{S}(q,\theta)r(t) + w(t)\,, \tag{11}$$

where $w(t)$ is an estimate of the noise signal. From the results in open loop identification $\hat{S}(q,\theta)$ can be identified consistently.

In the second step of the procedure the output relation is employed. It reconstructs the input signal $u^r(t)$ and identifies the close loop model using the reconstructed input and the measured output signal. Since $u^r(t)$ is available from the measurements through the command signal $r(t)$:

$$u^r(t) = \hat{S}(q,\theta)r(t) \tag{12}$$

and $\hat{P}(q,\theta)$ can be estimated in open loop way

$$y(t) = \hat{P}(q,\theta)u^r(t) + \hat{Q}(q,\theta)v(t)\,, \tag{13}$$

where $v(t)$ is the white noise signal, and $\hat{Q}(q,\theta)$ is an estimate of the transfer function between the noise and the output signal. A block diagram, indicating the recasting of the closed loop problem into two open loop problems, is sketched in *Fig. 3.*

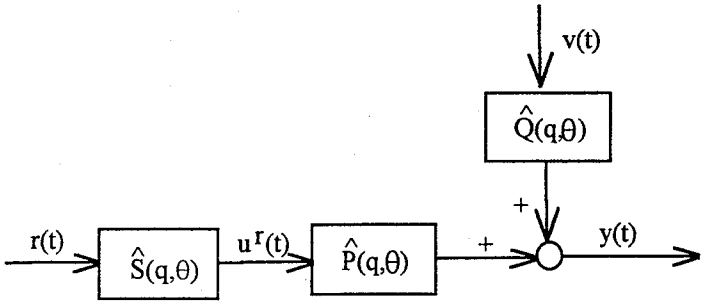


Fig. 3. Block diagram of the two-stage method

4.2. The Coprime Factorization Method

However, a relevant question is, specially under closed loop experimental conditions, how to identify systems that are unstable. The coprime factorization method can be applied for identification of an unstable closed loop system, too. It is introduced in ZHU and STOORVOGEL, 1992, and SCHRAMA, 1991.

Let us define $S(q)$ and $W(q)$ factors that are the common factors of the input ans output equations (1) and (2)

$$S(q) = C(q)[I + P(q)C(q)]^{-1}, \tag{14}$$
$$W(q) = [I + P(q)C(q)]^{-1}. \tag{15}$$

Applying these notations the following two equations can be obtained.

$$u(t) = S(q)r(t) - C(q)W(q)v(t), \tag{16}$$
$$y(t) = P(q)S(q)r(t) + W(q)v(t). \tag{17}$$

The essence of this method is that $S(q)$ can be identified from (16) based on the reference and the input signals independently from $P(q)S(q)$, which can be identified from (18) based on the reference and the output signals. If the reference signal $r(t)$ is uncorrelated with the disturbance $v(t)$, then these are open loop identification problems. The results of these identifications are $\hat{D}(q,\theta)$ and $\hat{N}(q,\theta)$ as it can be seen in (18) and (19). The two transfer functions between $r(t)$ and $u(t)$, $y(t)$ are identified according to the scheme of *Fig. 4*

$$u(t) = \hat{D}(q,\theta)r(t) + \hat{R}(q,\theta)v(t), \tag{18}$$
$$y(t) = \hat{N}(q,\theta)r(t) + \hat{Q}(q,\theta)v(t). \tag{19}$$

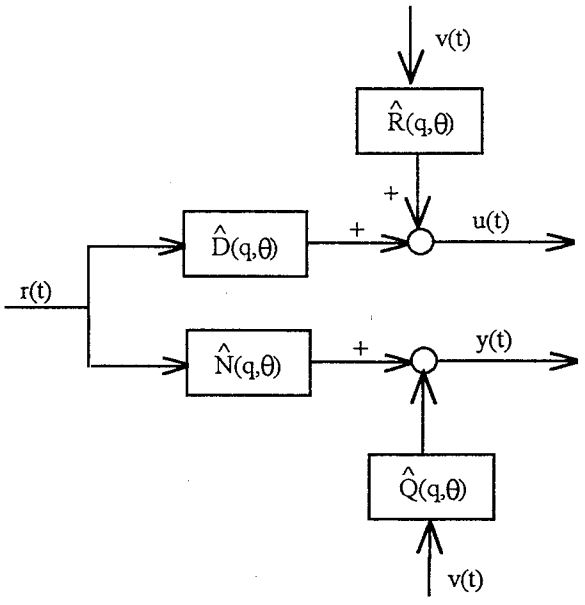


Fig. 4. Block diagram of the coprime factorization method

The pair $P(q)S(q)$ and $S(q)$ can be considered as a factorization of $P(q)$, since $P(q)S(q)[S(q)]^{-1} = P(q)$ assuming that $S(q)$ is non-singular, so the identified model can be obtained applying the identified $\hat{D}(q,\theta)$ and $\hat{N}(q,\theta)$

$$\hat{P}(q,\theta) = \hat{N}(q,\theta)[\hat{D}(q,\theta)]^{-1} . \tag{20}$$

Since the closed loop system is stable, the two separate factors composing this factorization are also stable, and they can be identified from data measured in the closed loop. If the estimates of both factors are obtained from independent data sequences, the $\hat{P}(q,\theta)$ estimate will be consistent in the case that both estimators $\hat{D}(q,\theta)$ and $\hat{N}(q,\theta)$ are consistent.

5. Case Study for Active Suspension System

Conventional passive suspensions are typically composed of coil springs and hydraulic dampers. To improve ride comfort, these passive elements must be set on the soft side, but to improve handling, springs and dampers must be made stiffer to damp wheel vibration and to reduce body rolling and pitching motions. These contrary purposes from the classic suspension design compromise. In order to obtain both excellent ride comfort and handling, the suspension characteristics must be changed dynamically according to the demands of the situation. In other words, active control of the suspension is required. The structure of a quarter car model active suspension has been investigated by HROVAT, 1990.

In order to investigate the benefit of active suspension systems, the following two-degree-of-freedom quarter-car model is applied. Let vehicle sprung and unsprung mass be denoted by m_f and m_t, tire stiffness and damping of sprung mass be denoted by s_f and k_f and tire stiffness of unsprung mass be denoted by s_t as it can be seen in the *Fig. 5*. Let $\ddot{z}_t$ be the acceleration of the m_t, $\ddot{z}_f$ the acceleration of the m_f, d stands for the road disturbance.

The differential equations of the model in *Fig. 5* can be formulated as

$$\begin{aligned} m_f\ddot{z}_f + k_f(\dot{z}_f - \dot{z}_t) + s_f(z_f - z_t) + u = 0 , \\ m_t\ddot{z}_t + s_t(z_t - d) - s_f(z_f - z_t) - k_f(\dot{z}_f - \dot{z}_t) - u = 0 , \end{aligned} \tag{21}$$

whose state space representation can be considered as

$$\begin{aligned} \dot{x} &= Ax + Bu + Gd , \\ y &= Cx + Du , \end{aligned} \tag{22}$$

where $x = (\, z_t \quad z_f \quad \dot{z}_t \quad \dot{z}_f \,)^T$ is the state vector, $y = \ddot{z}_f$ is the output, while

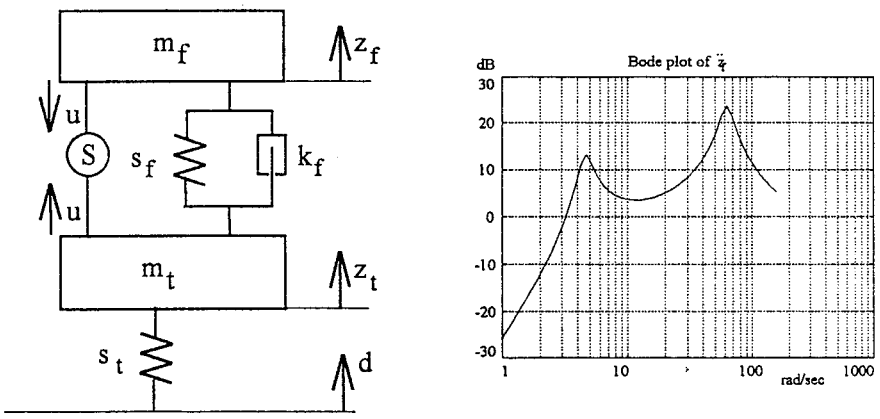


Fig. 5. Scheme of a quarter car model and Bode plot of $\ddot{z}_f$

the constant matrices are

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{s_f + s_t}{m_t} & \frac{s_f}{m_t} & -\frac{k_f}{m_t} & \frac{k_f}{m_t} \\ \frac{s_f}{m_f} & -\frac{s_f}{m_f} & \frac{k_f}{m_f} & -\frac{k_f}{m_f} \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{m_t} \\ -\frac{1}{m_f} \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ 0 \\ \frac{s_t}{m_t} \\ 0 \end{pmatrix},$$

$$C = \begin{pmatrix} -\frac{s_f + s_t}{m_t} & \frac{s_f}{m_t} & -\frac{k_f}{m_t} & \frac{k_f}{m_t} \end{pmatrix}, \quad D = \begin{pmatrix} \frac{1}{m_t} \end{pmatrix}.$$

Let the simulation parameters be $m_f = 400$ kg, $m_t = 33$ kg, $s_t = 9000$ N/m, $s_t = 120000$ N/m, $k_f = 500$ Ns/m. The controller is designed based on the identified model, and this controller is applied for the input u. The Bode plot of the open loop plant, i.e. of the passive suspension system, and one of the actual closed loop system, i.e. of the actual active suspension system can be seen in *Fig. 6.* The system has two peaks, the first is in 4.7 rad/sec, while the other is in 62.8 rad/sec.

The investigation of closed loop identification has been performed based on simulated data, which are demonstrated in *Fig. 7.* In the following the direct method, the two-stage method, and the coprime factorization method are compared. It can be performed since none of them assumes the knowledge of the controller, but they apply the measured signals.

Applying the **direct method** we have investigated, how the different order identified models approach the behavior of the actual plant. The left hand side of *Fig. 8* shows the estimation of the 4.7 rad/sec peak. It can be seen that only the higher models approach this peak. Moreover, the right hand side of *Fig. 8* shows the estimation of the 62.8 rad/sec peak with the same structures. It can be seen that all of them well fit to the peak. This is

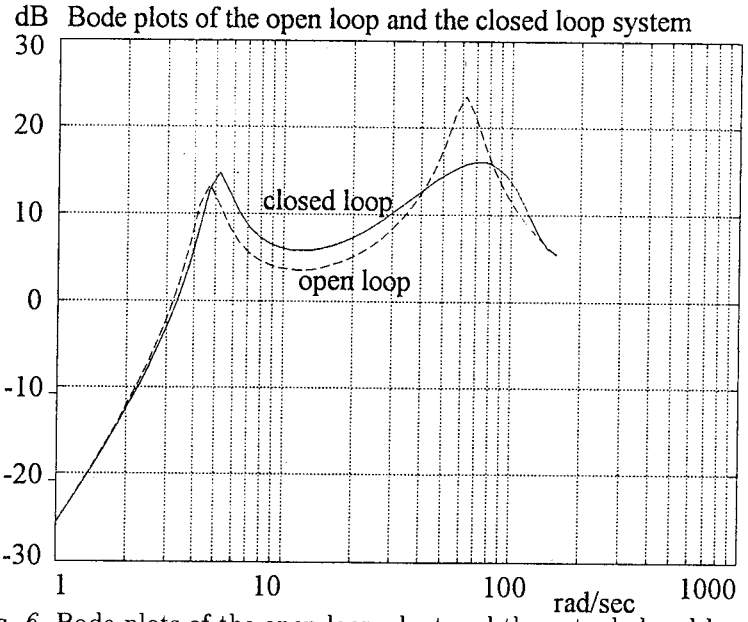


Fig. 6. Bode plots of the open loop plant and the actual closed loop system

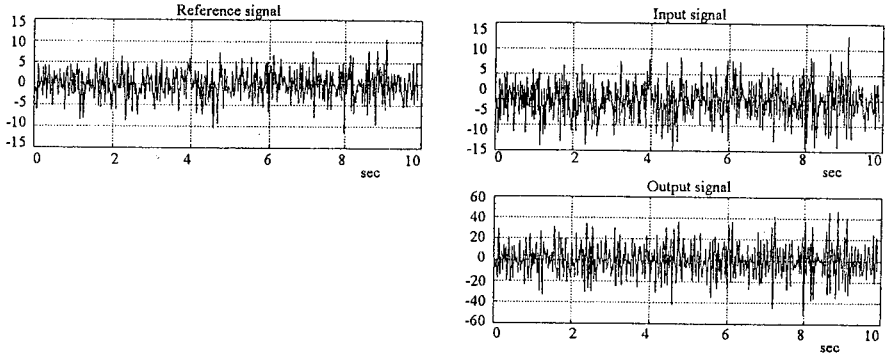


Fig. 7. Sampled time series based on closed loop system

why we select the structure where the autoregressive (AR) order is 30, and the input (INP) order is 30 as well, i.e. the model is ARX(30,30).

On the left hand side of the *Fig. 9* the AR and the INP parameters of the identified model are shown based on direct method as a function of lag. The frequency domain representation of the identified model approaches well the plot of the original systems as it can be seen on the right hand side of the *Fig. 9.*

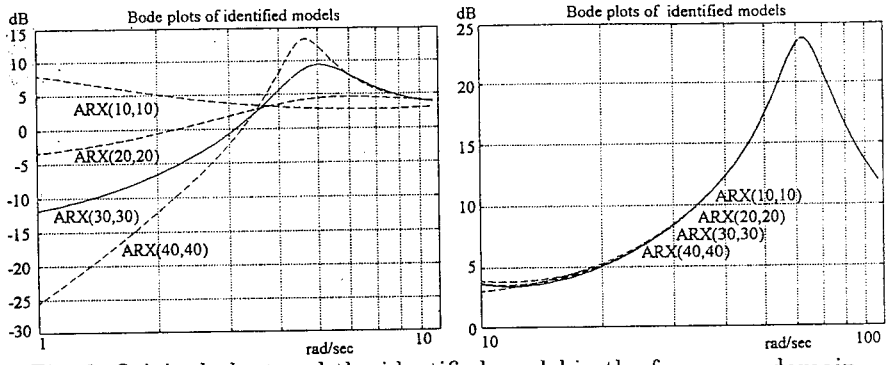


Fig. 8. Original plant and the identified model in the frequency domain

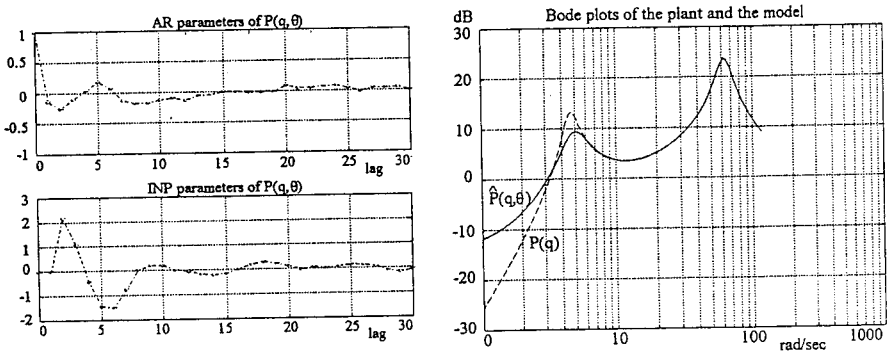


Fig. 9. Model parameters and Bode plot of the identified model based on direct method

The aim of the first step of the **two-stage method** is to reconstruct the input signal using the identified $\hat{S}(q,\theta)$ transfer function between $r(t)$ and $u(t)$ signals. The estimated parameters of the $\hat{S}(q,\theta)$ and its frequency domain representation can be seen in *Fig. 10.* The reconstructed input signal can be seen in *Fig. 11.*

In the second step of the two-stage method, $\hat{P}(q,\theta)$ model can be estimated in open loop way between the reconstructed input and the measure output signals. The order of the identified model is ARX(30,30). *Fig. 12* shows the estimated parameters as a function of lag, and illustrates its frequency domain representation.

Figs. 13 and *14* show the AR and the INP parameters and the frequency domain representation of the $\hat{D}(q,\theta)$ and the $\hat{N}(q,\theta)$ models in the **coprime factorization method**. Based on the $\hat{D}(q,\theta)$ and the $\hat{N}(q,\theta)$ the obtained $\hat{P}(q,\theta)$ model is ARX(59,59). The frequency domain representation of the calculated model can be seen in *Fig. 15.*

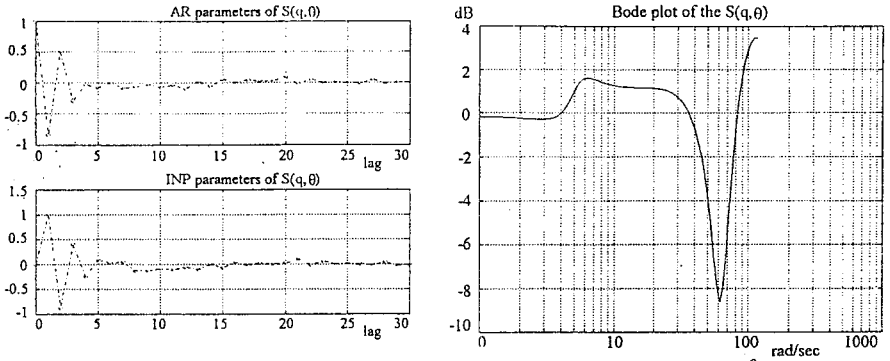


Fig. 10. Model parameters and Bode plot of the identified $\hat{S}(q,\theta)$

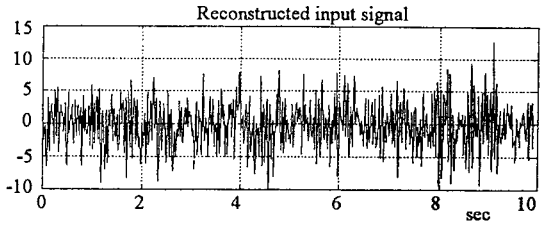


Fig. 11. The reconstructed input signal

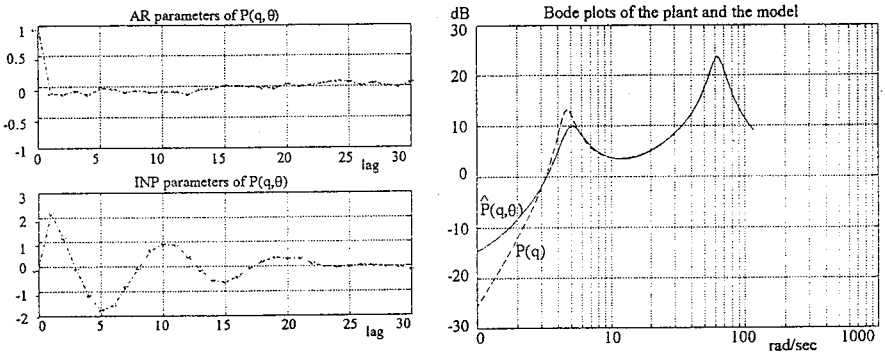


Fig. 12. Model parameters and Bode plot of the model based on two-stage method

In this example we have experienced that the classical direct method supplies the least dimension model in comparison with the other two approaches. The two-stage method has given higher order dimension in this example, but the order of the estimated model can be reduced by improving the reconstruction of the input signal. This method also performs in open loop way, but it takes into account the feedback effect. The coprime factorization method has resulted the highest order model, that can be computed

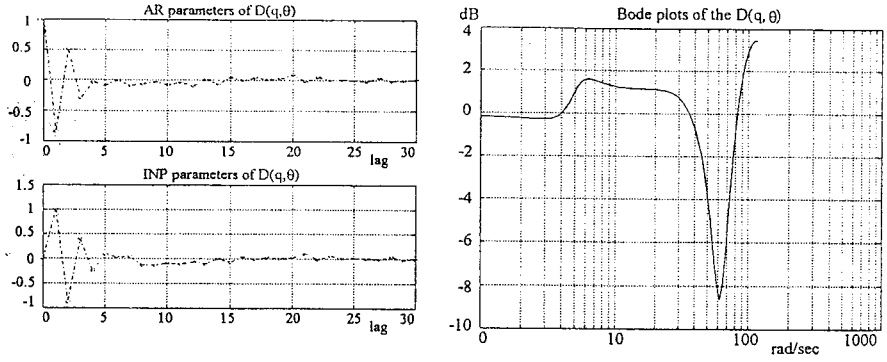


Fig. 13. Model parameters and Bode plot of the identified $\hat{D}(q,\theta)$

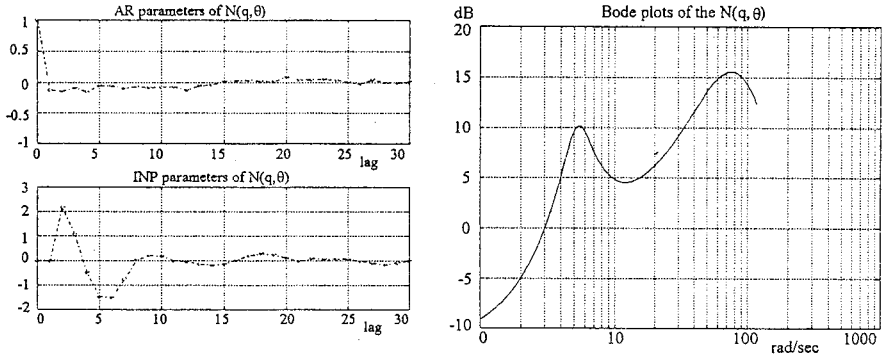


Fig. 14. Model parameters and Bode plot of the identified $\hat{N}(q,\theta)$

from the identified $N(q)$ and $D(q)$ factors. We can get lower order structure of the identified model if $N(q)$ and $D(q)$ have lower degree. The last method is able to handle unstable closed loop systems, too, which is impossible for the two other methods.

6. Conclusion

In this paper we have highlighted that closed loop identification has to be performed for closed loop design. We have summarized the most important closed loop identification methods (traditional direct/indirect, the two-stages and the coprime factorization methods) with respect to the active suspension design example. We have concluded that these methods are well applicable and can be easily used.

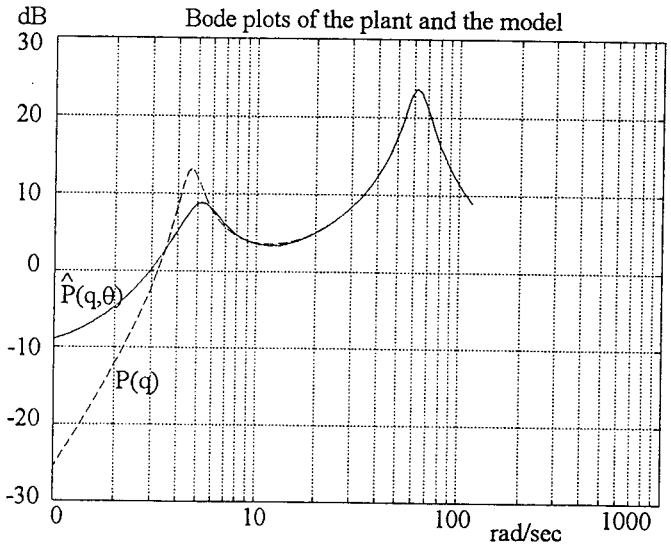


Fig. 15. Bode plot of the identified model based on the coprime factorization method

Acknowledgement

This project has been supported by the Hungarian National Foundation (OTKA) through grant No. T-016418 which is gratefully acknowledged.

References

[1] GEVERS, M.: Connecting Identification and Robust Control: A New Challenge, *Proceedings of the 9th IFAC/IFORS Symposium*, Budapest, pp. 1–10, 1991.

[2] HANSEN, F. – FRANKLIN, G. – KOSUT, R.: Closed-Loop Identification Via the Fractional Representation Experiment Design, *Proceedings of the American Control Conf.*, Pittsburgh, pp. 1422–1427, 1989.

[3] HROVAT, D.: Optimal Active Suspension Structure for Quarter-Car Vehicle Models, *Automatica*, Vol. 26, No. 5, pp. 845–860, 1990.

[4] LJUNG, L.: System Identification: Theory for the User, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1987.

[5] SCHRAMA, R. J. P.: An Open-Loop Solution to the Approximate Closed-Loop Identification Problem, *Proceedings of the 9th IFAC/IFORS Symposium*, Budapest, pp. 1602–1607, 1991.

[6] SKELTON, R. E.: Model Error Concepts in Control Design, *International Journal of Control*, Vol. 49, No. 5, pp. 1725–1753, 1989.

[7] SÖDERSTRÖM, T. – LJUNG, L. – GUSTAVSSON, I.: Identifiability Conditions for Linear Multivariable Systems Operating Under Feedback, *IEEE Trans. Automatic Control*, Vol. 21, pp. 837–840, 1976.

[8] Van den HOF, P. M. J. – SCHRAMA, R. J. P.: An Indirect Method for Transfer Function Estimation from Closed Loop Data, *Proceedings of the 31st Conf. on Decision and Control*, Tucson, pp. 1702–1706, 1992.
[9] ZHU, Y. C. – STOORVOGEL, A. A.: Closed Loop Identification of Coprime Factors, *Proceedings of the 31st Conf. on Decision and Control*, Tucson, pp. 453–454, 1992.

PERIODICA POLYTECHNICA SER. TRANSP. ENG. VOL. 24, NO. 2, PP. 95-101 (1996)

IDENTIFICATION OF SOME MOTOR-CYCLE COMPONENTS RELIABILITY

Miroslav KOPECKY

Department of Mechanics and Strength
University of Transport and Communications
Zilina, Slovakia

Received: Nov. 7, 1994

Abstract

Modern technology has advanced the development of many new materials and products which in turn have created the need for new and advanced test methods. The material must be thoroughly investigated for mechanical properties such as fatigue life or maximum strength to permit efficiency and economy, as well as reliability and safe design of component structures. The method described in this paper deals with the stress distribution and application of statistical longevity of some motor-cycles components under operating the random loads. The results of its application are restricted to load-carrying motor-cycles components of small capacity.

Keywords: fatigue, statistical longevity, probability density function failure, reliability.

1. Introduction

Acquisition of the command with two-wheeled road vehicle for laboratory loading system contains identifications of the dynamic characteristic features in operating state, during random loads. The usual way to determine the equivalent loading of a random evolution is displayed in *Fig. 1.*

The signal of response may be analysed by statistical characteristic of stochastic function.

The random excitation of two-wheeled road vehicle is restricted to the random change of road surface undulations which provides a vertical input displacement into the tire of the vehicle.

The laboratory test is to simulate the dynamic response of the vehicle by subjecting each wheel at random varying displacement imposed by means of a simulator as shown in *Fig. 2.*

A motorcycle running along a road is subjected to two vertically imposed displacements, one at each wheel. The description of the road surface must be complete enough to describe adequately the displacement imposed at each wheel at least in statistical terms and the correlation between the two displacements.

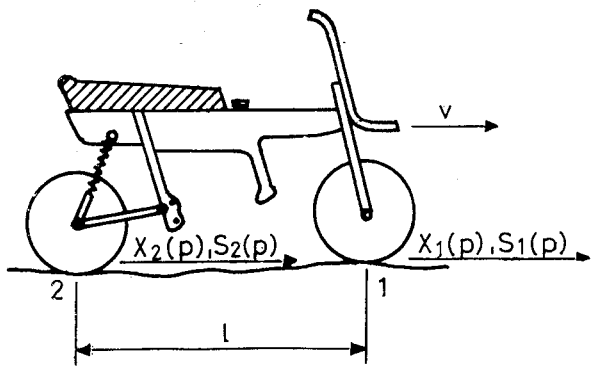


Fig. 1. Layout of inputs for a two-wheeled vehicle

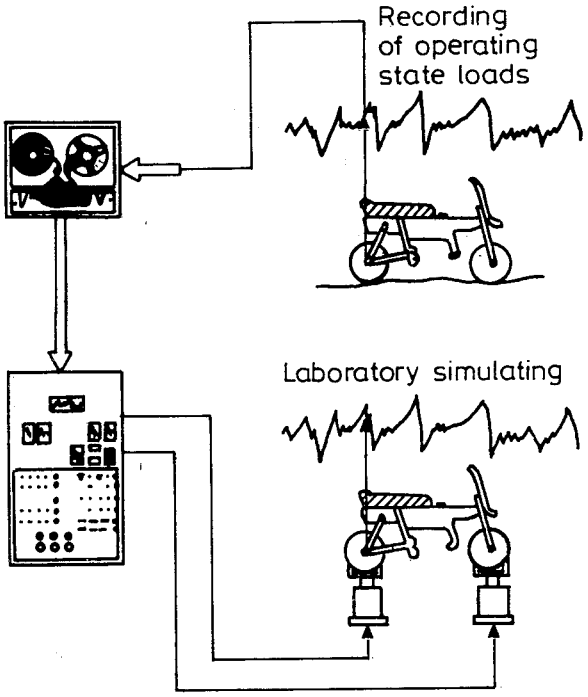


Fig. 2. Layout of laboratory test

2. Application of Statistical Longevity

The starting point of theoretical solutions reliability is the WEIBULL model [3]. The dependence between random loads and life, $\mathbf{N}_f$, of components must be completed by a variable, $\mathbf{R}(N_f)$, which expresses digital quarantee in the probability form.

A three-parameters distribution may be expressed as

$$\mathbf{R}(N_f) = \exp(-((N_f - N_{\min})/(N_{\text{sig}} - N_{\min}))^k) \tag{1}$$

where: $N_{\min}$ is a minimum of the longevity
N_{sig} is a modal value of the longevity
k is a parameter of distribution.

The probability density function related to (1) is of the form

$$f(N_f) = \frac{k}{N_{\text{sig}} - N_{\min}} \cdot \left[\frac{N_f - N_{\min}}{N_{\text{sig}} - N_{\min}}\right]^{k-1} \cdot \exp\left(-\left(\frac{N_f - N_{\min}}{N_{\text{sig}} - N_{\min}}\right)^k\right) . \tag{2}$$

The determination of the parameters of this distributions $\mathbf{k}, \mathbf{N}_{\text{sig}}, \mathbf{N}_{\min}$ are achieved by the moment of function (2) numerically.

Common value of n-th moment for variables

$$\frac{N_f - N_{\min}}{N_{\text{sig}} - N_{\min}} \quad \text{is:} \quad \mathbf{m}_n = \Gamma\left(1 + \frac{n}{k}\right)$$

The first and second central moment of basic distribution *Eq.* (2) are:

$$\mathbf{m}_1 = \Gamma\left(1 + \frac{1}{k}\right) \; ; \quad \mathbf{m}_2 = \Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right) .$$

The coefficient of obliquity is determines with second and third central of moment:

$$\frac{\mathbf{m}_3}{\mathbf{m}_2^{3/2}} = \frac{\left[\Gamma\left(1 + \frac{3}{k}\right) - 3\Gamma\left(1 + \frac{2}{k}\right) \cdot \Gamma\left(1 + \frac{1}{k}\right) + 2\Gamma^3\left(1 + \frac{1}{k}\right)\right]}{\left[\left(\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right)^{3/2}\right]} . \tag{3}$$

If we return to original variable $\mathbf{N}_f$, the first moment will be:

$$\mathbf{m}_1(N_f) = N_{\min} + (N_{\text{sig}} - N_{\min})^{1/k} \cdot \Gamma\left(1 + \frac{1}{k}\right) \tag{4}$$

and it is a reply to estimate the moment $\mathbf{m}_1(N_1)$ for basic random selection:

$$\mathbf{m}_1(N_1) \approx \mathbf{N}_S \tag{5}$$

where: $\mathbf{N}_S$ is a middle value of longevity.

The dispersion of original variable $\mathbf{N}_f$ may be expressed in a term:

$$\mathbf{m}_2(N_f) = (N_{\text{sig}} - N_{\min})^{2/k} \cdot \left(\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right) \tag{6}$$

and the estimate of moment is

$$\mathbf{m}_2(N_f) \approx \mathbf{S}_N^2 , \tag{7}$$

where: $\mathbf{S}_n$ is the standard deviation.

From *Eqs.* (4) and (5) is:

$$\mathbf{N}_{\text{sig}} = N_S + \mathbf{S}_N \cdot A(k) . \tag{8}$$

From *Eqs.* (6), (7) and (8) is:

$$\mathbf{N}_{\min} = N_S - S_N \cdot D(k) . \tag{9}$$

The application functions in *Eqs.* (8) and (9) are:

$$\mathbf{A}(k) = \frac{\left[1 - \Gamma\left(1 + \frac{1}{k}\right)\right]}{\left[\left(\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right)^{1/2}\right]}$$

and

$$\mathbf{D}(k) = \frac{\left[\Gamma\left(1 + \frac{1}{k}\right)\right]}{\left[\left(\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)\right)^{1/2}\right]} .$$

With parameters of distribution, we may define the result by the statistical curve of longevity, which in a form of probability characterized the longevity form *Eq.* (1)

$$\ln(-\ln \mathbf{R}(N_f)) = k(\ln(N_f - N_{\min}) - \ln(N_{\text{sig}} - N_{\min})) .$$

But for the case $N_{\min} = 0$, the function of probability of longevity, *Eq.* (1), will be reduced to two-parameters of term:

$$\mathbf{R}(N_f) = \exp\left(-\left(\frac{N_f}{N_{\text{sig}}}\right)^k\right) . \tag{10}$$

The probability density function related to *Eq.* (10) is:

$$\mathbf{f}(N_f) = \frac{k}{N_{\text{sig}}} \cdot \left(\frac{N_f}{N_{\text{sig}}}\right)^{k-1} \cdot \exp\left(-\left(\frac{N_f}{N_{\text{sig}}}\right)^k\right) . \tag{11}$$

Estimates of parameters k, $\mathbf{N}_{\text{sig}}$ by characteristic value $\mathbf{N}_S$, $\mathbf{S}_N$ may be expressed:

from *Eqs.* (4) and (5)

$$\mathbf{N}_{\text{sig}} = \frac{N_S}{C(k)}$$

from *Eqs.* (6) and (7)

$$\mathbf{N}_S = S_N \cdot D(k)$$

where: $\mathbf{C}(k) = \Gamma\left(1 + \frac{1}{k}\right)$.

The functions $\mathbf{A}(k)$, $\mathbf{C}(k)$, $\mathbf{D}(k)$ and $\mathbf{B}(k)$ in *Eq.* (3), are introduced for the practical application in *Eq.* (1) for variables $\frac{1}{\mathbf{k}}$.

3. Experiment and Results

The applications of this method in this paper are restricted to load-carrying parts of motor-cycles of small capacity. Tests are frequently completed on construction subassemblies, such as they are shown in *Fig. 3.*

The results of laboratory simulating test for a frame construction are shown for illustration. The frame is the most important load-carrying part of a motor-cycle. The test of simulation regime for a frame construction has been made upon the special purpose machine in a laboratory.

The laboratory test has been made for 3 alternative frame contructions. Experimental results of the laboratory test are shown in *Table 1.*

Alternative **A** is seen to be the most reliable. The results of random excitation in operating state are to be seen in *Table 2.*

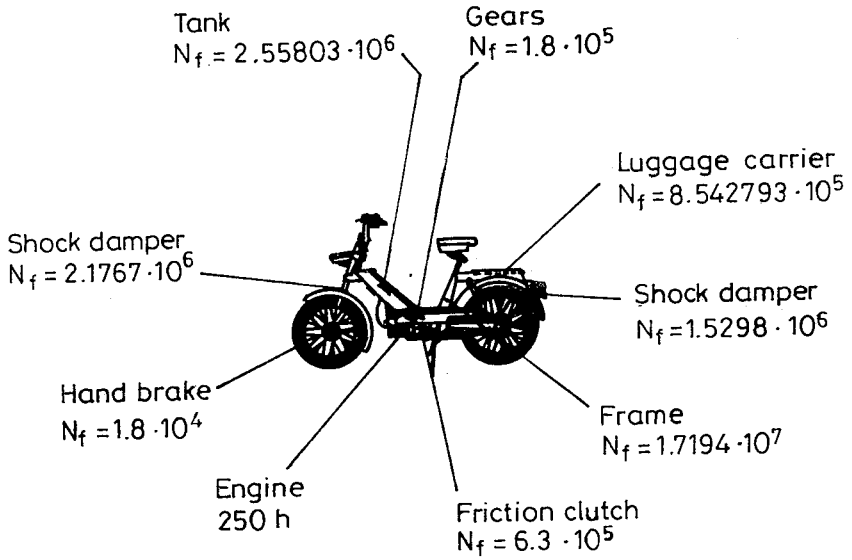


Fig. 3. View of the load-carrying parts of motor-cycles

The statistical curve of longevity for **A** alternative frame construction from *Eq.* (10) is shown in *Fig. 4.*

Table 1. Frame constructions

Reliability: Number of cycles to failure	Alternative: A:	B: $\times 10^6$	C:
N_{f1}	22.4656	6.8759	12.1511
N_{f2}	25.4133	1.2145	3.8624
N_{f3}	22.4967	4.4231	1.8765
N_{f4}	48.2970	4.0824	2.6661

Table 2. Frame constructions

Statistical moment	Parameters of distributions			$\mathbf{R}(N_f)$ %
$\mathbf{m}_1 = 992.36$	$\mathbf{k}$	=	1.062	
$\mathbf{m}_2 = 1690.24$	$\mathbf{N}_{sig}$	=	$5.25701 \cdot 10^6$	
$\mathbf{m}_3 = 1346291.7$	$\mathbf{N}_{min}$	=	$8.82048 \cdot 10^4$	
	$\mathbf{N}_f$	=	$8.542793 \cdot 10^5$	0.9999

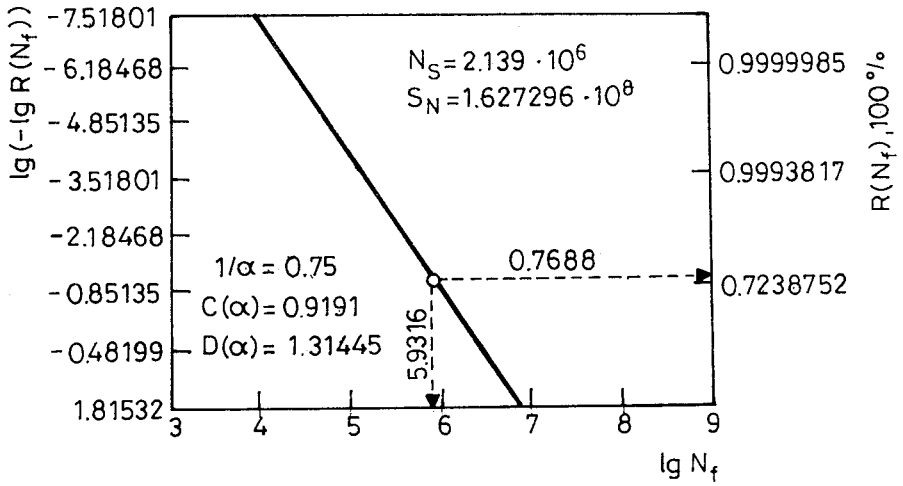


Fig. 4. Statistical curve of the longevity

Eq. (2) is an answer to the curves of probability density, from *Fig. 5* in our case for frames.

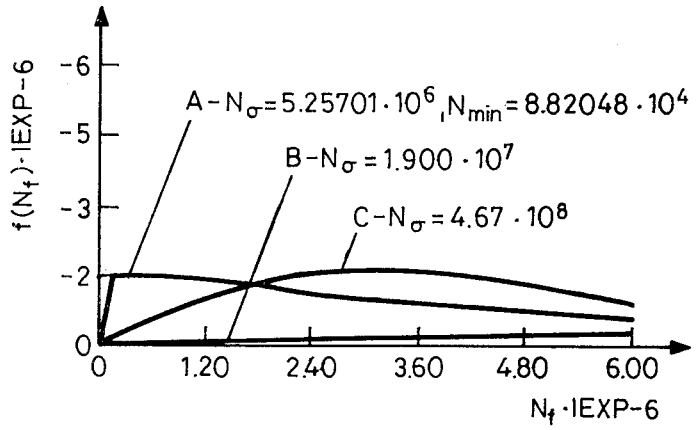


Fig. 5. Curves of probability density distribution

4. Conclusion

The test of the longevity of load-carrying parts of motor-cycles at laboratory makes variable extreme conditions of random excitation in operating state possible.

The application of this method shortens knowledge of the time to failure of machine components for transportation and contributes to the safety and economy of mechanical systems.

References

[1] KOPECKY, M. (1990): Experimental-Numerical Method of Random Loading Analysis. In: *9th Inter. Conf. on Experimental Mechanics*, Vol. 3, pp. 1006–1012, Copenhagen.

[2] KOPECKY, M. (1993): Life Assessment of Some Load-Carrying Machine Components for Transportation, In: *Structural Safety*, Vol. 12, pp. 145–149, Elsevier Science Publishers, The Netherlands.

[3] WEIBULL, W. (1951): A Statistical Distribution Function of Wide Applicability, *Journal of App. Mech*, No. 3.

PERIODICA POLYTECHNICA SER. TRANSP. ENG. VOL. 24, NO. 2, PP. 103-116 (1996)

THE DYNAMIC RESPONSE OF AIRCRAFT WHEEL

Lajos KISS
Szolnok College of Flight Officers
H-5008 Szolnok P.O.B.1, Hungary

Received: November 30, 1994

Abstract

The author of the lecture has collected and developed a method for analyzing the characteristics of parametes affecting aircraft tyre control forces, prediction of aircraft braking friction on runways, dynamics tyre-soil contact surface interaction model for aircraft ground operation and the dynamic response of an aircraft wheel to variation in runway friction.

Keywords: aircraft, dynamic response, braking friction, tyre model, surface.

1. Introduction

The author of the lecture has collected and developed a method for analyzing the characteristics of: parameters affecting aircraft tyre control forces, prediction of aircraft braking friction on runways, dynamic tyre–soil contact surface interaction model for aircraft ground operation and the dynamic response of an aircraft wheel to variations in runway friction.

2. Parameters Affecting Aircraft Tyre Control Forces

The movement of an aircraft is primarily the result of frictional force between the tyre, and the ground surface. As a result of the bad condition or the snowy, icy surface of the runway, or the transversal inclination of highways during the rolling of the tyres the probability of their transversal motion (crawling) is increased.

About the F_z force, resulting from the sideways sliding, we can get answers from:

- the F_n load on one wheel
- the lateral sliding angle β [1], [3]. (*Fig. 1*)

This relationship can be further modified by the fact that we can alter the tyre's rigidness by reducing the tyre pressure. This is permitted to increase off-road capability. Similarly, the quality of the runway and the tyre-compound used is also having effect on the above described problem [4].

The behaviour of the tyre under the effect of the moment of braking is determined by the rigidness and the frictional parameters of the tyre. We can see (on *Fig. 2*) the rolling of a wheel at the speed of V_x under the load of F_n.

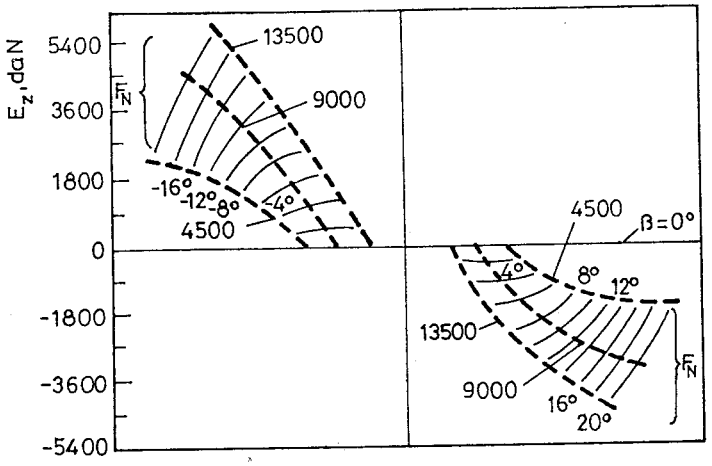


Fig. 1. F_z, F_n and β relationship

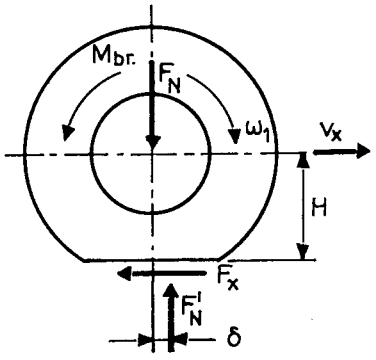


Fig. 2. The wheel model

Form this, the relative turn-over can be calculated, which is:

$$S = \frac{\omega_0 - \omega_1}{\omega_0} . \tag{1}$$

The $S = 0$ means the free-rolling of the wheel while $S = 1$ means that the wheel is blocked [1].

Under braking, the speed of motion (*Fig. 3*) and the amount of dirt on the runway (*Fig. 4*) are the major influencing factors on the friction of

the wheel. On *Fig. 3* at the point $S = 0$ the C_s rigidness – characteristic of every tyre – can be marked. On *Fig. 5* the effect of the thickness of the water layer on a wet runway is shown [3].

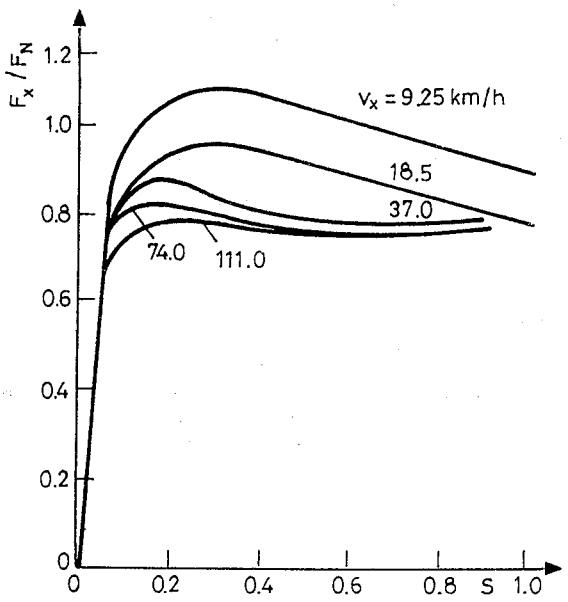


Fig. 3. Influence of the speed on the friction force

The thing of interest about the diagram is that the points of stable and unstable operation can be marked (*Fig. 6*) [4].

In practice, the modern anti-blocking systems (ABS) adjust the maximum coefficient of friction between the tyre and the runway. On wet, dirty or improperly prepared runways, this marginal relationship suddenly decreases which results in the loss of control in the steering of the aircraft along the runway or in the loss of stability during landing. Hence, this reduces the usability of maneuver airfields and highways. This is shown on *Fig. 7* [4].

This gives rise to the problem that for the operators of the aircraft, it is important to know better the most optimal relationship achievable between the longitudinal and transverse forces without the worsening of the parameters of the takeoff and landing.

3. Prediction of Aircraft Braking Friction on Runways

During the researches made by the US Air Force, NASA, and the FAA between 1968 and 1980, it was shown that this problem cannot easily be

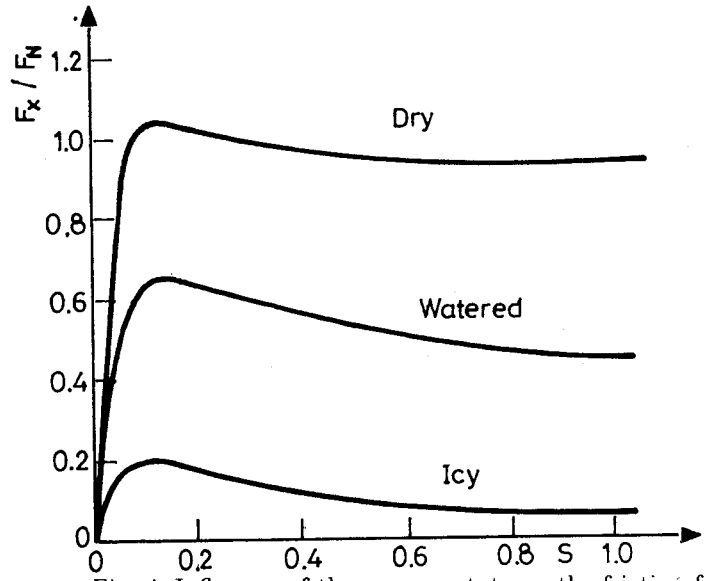


Fig. 4. Influence of the runway state on the friction force

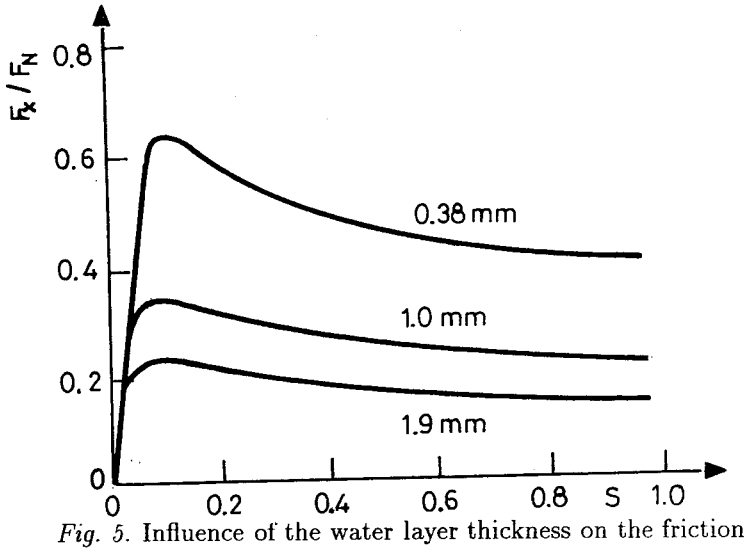


Fig. 5. Influence of the water layer thickness on the friction force

modelled. From their experiments they have reduced that there are 47 such characteristic parameters of the tyre which need to be examined for the correct description of the rolling tyre's frictional relationship.

I think, in harmony with part I, the following 19 basic parameters should be considered:

parameters of the tyre:

1 load F_n, F_z, F_x
2 inside pressure $P_{t,0}, P_{t;1}$
3 construction of the tyre (e.g. radial, diagonal, etc.)
4 texture of the abrasive layer
5 size of the wheel (D^*B, d^*b)
6 the method of the strengthening of the abrasive layer
7 the type of material used for the strengthening of the abrasive layer (natural or synthetic)
8 the amount of wear of the tyre

parameters of the liquid layer covering the runway

9 viscosity
10 density
11 thickness of the layer

parameters of the runway surface

12 microstructure
13 macrostructure
14 friction measured with a polished tyre
15 friction measured with a worn, eroded tyre
16 stability against erosion
17 temperature
18 the marginal point between rolling and blocking
19 frictional coefficient under braking

In other words, these parameters include all the parameters of grass airfields, temporary covering layers, concrete runways in bad condition and of highways together with the parameters of the aircraft tyre.

For example the sliding theory developed by the NASA is interesting because it gives explanation for the rolling of the tyre on wet surfaces but it is inaccurate about the calculation of the forces creating this relationship [2].

The experiments conducted by the NASA have confirmed the relationship between some parameters on wet runways:

- radial load
- size, structure, texture of the tyre
- depth of canals on the tyre surface
- composure of the abrasive layer
- temperature
- thickness of water layer
- structure of the top (covering) layer of the runway
- mode and position of the turning wheel

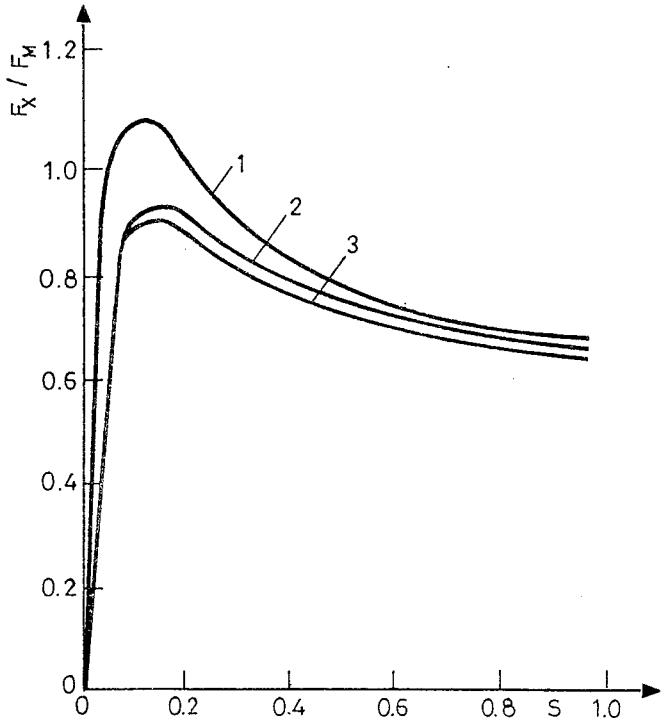


Fig. 6. Stable and unstable operation points

Fig. 8 shows the effects of the wear of the tyre of a 32*8, 8 VII wheel having longitudinal canals and an inside pressure of 1.03 MPa, having been used on a wet runway [1], [2].

Markings of the diagram are:

1. average μ
2. rolling speed
3. wear of tyre in percentages

4. Dynamic Model for the Interaction between Tyre and Soil during Aircraft Ground Operation

To ensure the operation of aircraft from grass airfields – along other considerations – it must have certain 'cross-country' capability [4].

The 'ability-to-pass' include the following most important conditions:

- the determination of the effects of the deformation of the soil
- the condition of coming into motion from standstill

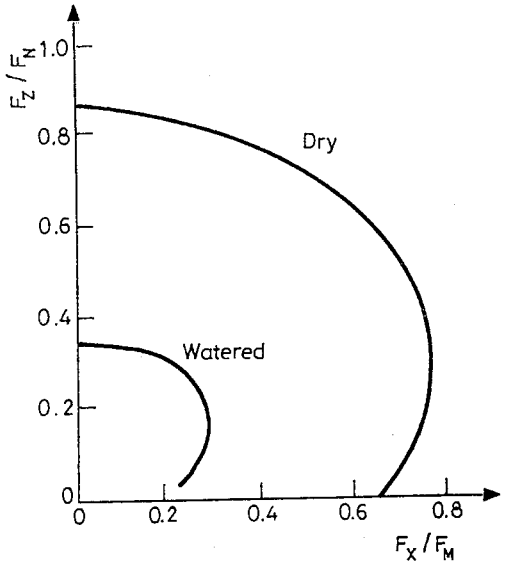


Fig. 7. F_x/F_n and F_z/F_n relationship

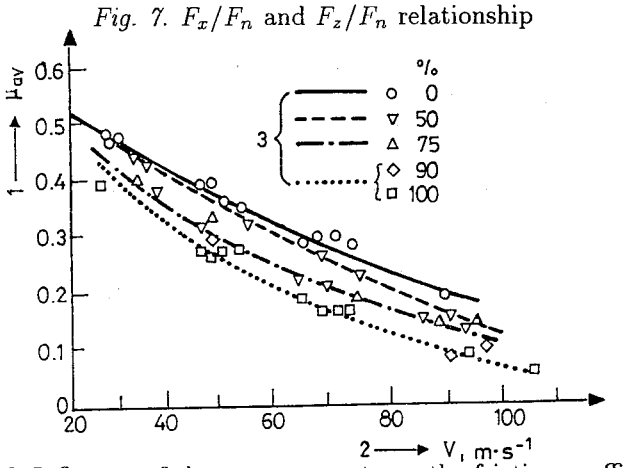


Fig. 8. Influence of the tyre wear rate on the friction coefficient

- the determination of the allowable imprint of the tyre
- the ensuring of the necessary length of runway for a takeoff-run
- the landing forces of the undercarriage

I suppose that the 'ability-to-pass' can be increased by decreasing the tyre pressure from $P_{t,0}$ to a smaller value of $P_{t,l}$ although this alters the tyre's rigidness as well.

The coefficient of the resistance of the soil deformation can be deter-

mined with the newly altered rigidness:

$$f_1 = \frac{q_{ff}}{c \cdot \xi \cdot \alpha_{\text{soil}} \cdot k_1}, \tag{2}$$

where:

q_{ff} : the specific load on the main wheels;
c : the ratio of the depressed surfaces (*Fig. 9*);
ξ : the correctional factor due to the change in the tyre's rigidness, (*Fig. 10*);
k_1 : a factor which in the case of several wheels being on the main-wheel strut takes into consideration the effect of the second wheel on the deformation of the soil.

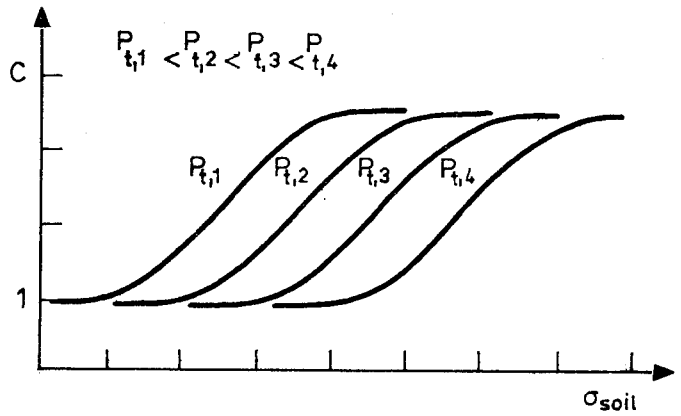


Fig. 9. The depressed surface ratio

In such cases, the necessary thrust of the aircraft to start from standstill is:

$$\mu_p = \frac{F_p}{G} \geq k_{\text{soil}} \cdot f_1, \tag{3}$$

where

F_p : the necessary thrust of the powerplant, N;
μ_p : the thrust-to-weight ratio
G : weight of the aircraft
k_{soil} : correcting factor, which takes into consideration 5 minutes of standing before starting on different soils of different state.

Thanks to the large thrust-to-weight ratio of modern aircraft it is not this previous data which restricts the operation of aircraft from grass airfields. The allowable maximum sinking of the tyre is:

$$h_{\text{allowable}} = 0.07 \cdot B \cdot D^{0.25}, \tag{4}$$

where B, D are the geometrical sizes of the wheel.

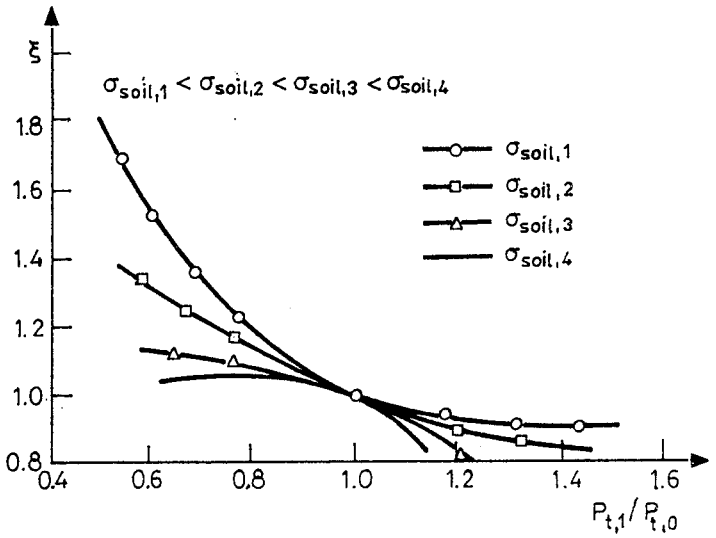


Fig. 10. The correctional factor-tyre's rigidness ratio relationship

Regarding the fourth criterion on the usability, it must be guaranteed that the takeoff and landing run of the aircraft must be smaller than the available length of grass runway together with the final security zone. These criteria may also restrict the utilization of grass airfields.

Amongst the conditions of operation, apart from the takeoff weight, the inside tyre pressure is of great importance according to BREWER's theorem. The permitted change in tyre pressure during takeoff and landing will be the following [1] [4]

$$p_{t,1} = (p_{t,0-cat}) \cdot \frac{F_{n_{t,0}}}{F_{n_{t,0-cat}}}, \tag{5}$$

$$p_{t,2} = (p_{t,0-cat}) \cdot \frac{F_{n_{land}}}{F_{n_{land-cat}}}. \tag{6}$$

The critical takeoff and landing speeds are:

$$v_{to.cr} = [v_{to.max-cat} + 20 \cdot (p_{t,1} - p_{t,0})\} \cdot \frac{F_{n_{to-cat}}}{F_{n_{to}}} \cdot \frac{p_{t,1}}{p_{t,0}}, \tag{7}$$

$$v_{land_{cr}} = [v_{land_{max-cat}} + 20 \cdot (p_{t,2} - p_{t,0})\} \cdot \frac{F_{n_{land-cat}}}{F_{n_{land}}} \cdot \frac{p_{t,2}}{p_{t,0}}, \tag{8}$$

where:

$p_{t,1}$	:	the minimum permitted pressure during takeoff, MPa;
$p_{t,2}$	:	the minimum permitted pressure during landing, MPa;
$F_{n.to}, F_{n.land}$	:	the actual load on the wheel during takeoff and landing, N;
$F_{n.to.cat}, F_{n.land.cat}$	:	the maximum load on the wheel according to catalogue, N;
$v_{(to.max.cat.)}, v_{(land.max.cat)}$	:	the manufacturer's restriction on speed after the installation of a type, according to catalogue, m/s.

The maximum weight-bearing capacity related to the above mentioned depending on the sub-layers of soil can be determined with the aid of the – locally well-known – Dorni method.

Obviously there are more modern methods of determining the 'ability-to-pass', but considering the tools at my disposal, I am able to calculate with this one.

In this part, the problem is caused by the fact that by reduction of the tyre pressure, its rate of exhaustion increases, its lifetime decreases [4].

5. The Dynamic Response of an Aircraft Wheel to the Variation in Runway Friction

At the Department of Aerospace at the Bristol University, they have been examining the problem of the ground motion of aircraft since 1970. For this research they have built a linear dynamometric device which was the first capable of examining the friction of an aircraft wheel when it was rolling on a softer surface than that of the tyre at a certain inside pressure. Later on, they have improved on this device by making it capable of examining the dynamic reactions of a braked aircraft on different solid-surfaced runways [5].

In doct. univers dissertation I have collected some temporary types of surfacing materials, which are used for the covering of maneuver airfields. Because the researches made at Bristol also include such materials, hence the dynamic properties of the wheel can be well-examined for the research of the earlier mentioned problem.

The experiment was done on the following types of surfaces:

1. WA : wet aluminium sheet
2. LWA : slightly wet aluminium sheet
3. DA : dry aluminium sheet
4. DS : dry aluminium sheet covered with sand grains

The appropriate sign for the type of surface used in the experiment can

be found on the time oscillograph. During the movement of the chassis across the different surfaces, the dynamic reactions of the wheel changes according to the frictional relationship and its value can be measured. Similarly, a detector is used on the surfaces which records the time when the wheel crosses the boundaries of the different sectors of surfaces

The results of a typical run can be seen on *Fig. 11* which has the marking of:

(1) markings of the boundaries of different sectors
(2) serial number of the type of surface
(3) time in ms

At the beginning, the wheel with a given vertical load and with a certain braking moment is crossing the DA surface which has a small coefficient of friction, which resulted in the blocking of the wheel (the relative turn-over was 1).

After this came along the DS sheet with a large coefficient of friction. The increasing frictional force starts to rotate the wheel, thus the relative turn-over is 0. The process is determined by the frictional coefficient μ, F_s frictional force movement, ω_w angular speed of the wheel, and the S_ω relative turn-over S_w [5].

The experiments call our attention to two phenomena, which are based on the flexibility of the tyre. The first one is apparent when the primarily blocked wheel works along the surface of a large coefficient of friction, the increase in μ is halted while the wheel reacts and starts to slow down. The second phenomenon can be determined from expriments. After the intensive turn-over, on the DS surface, quickly dampening small amplitude angular velocity oscillations of around 62 Hz can be seen. This phenomenon results from the fact that there is a relative motion between the wheel-base and the surface of contact.

Experiment can be shown with mathematical methods as well. The two-free-axis model is shown on *Fig. 12.*

The equations of motion are:

$$m \cdot a = -F_s \,, \tag{9}$$

$$I_1 \cdot \varepsilon_W = (\omega_t - \omega_W) \cdot K_t + (\alpha_t - \alpha_W) \cdot C_t - M_{br} \,, \tag{10}$$

$$I_2 \cdot \varepsilon_t = H \cdot F_s - (\omega_t - \omega_W) \cdot K_t - (\alpha_t - \alpha_W) \cdot C_t - F_n \cdot \left(\mu_g \cdot H - \frac{F_s}{C_s}\right) \,, \tag{11}$$

$$F_s = \mu \cdot F_n \,, \tag{12}$$

$$S_t = 1 - r_{e,0} \cdot \frac{\alpha_t}{x} \,, \tag{13}$$

$$S_w = 1 - r_{e,0} \cdot \frac{\alpha_W}{x} \,, \tag{14}$$

$$r_{e,0} = r_0 - \frac{\delta_0}{3} \,, \tag{15}$$

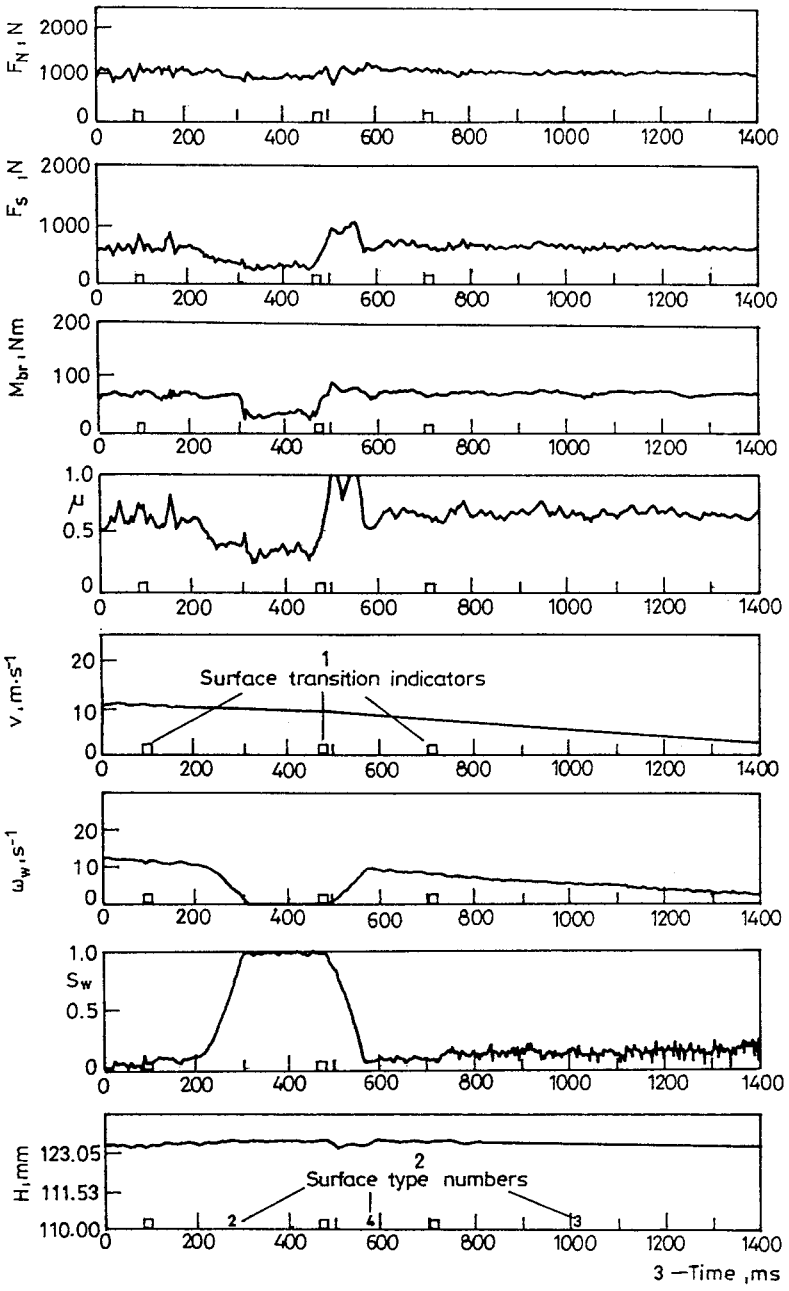


Fig. 11. Typical run's results

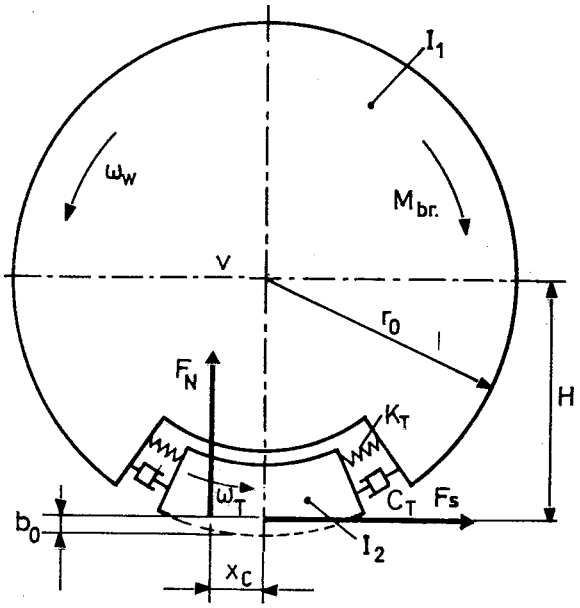


Fig. 12. The two-free-axis model

where K_t, C_t and I_z were determined through experiments. The coefficient of friction μ was determined through experiments as well.

When the relative turn-over is positive, that is

$$\mu = \mu(S_t, P_{t,0}, x, M_{brake}, \text{type of surface}); \quad S_t > 0 \tag{16}$$

when $S_t = 0$, then:

$$\alpha_t = \frac{x}{r_{e,0}}; \quad \omega_t = \frac{v}{r_{e,0}}; \quad \varepsilon_t = \frac{a}{r_{e,0}}. \tag{17}$$

We can diminish the variable ω_t and its derivatives from the equation. Therefore with the combination of the equations we get the following:

$$m = \frac{\left(\frac{x}{r_{e,0}} - a_w\right) \cdot K_t + \left(\frac{v}{r_{e,0}} - w_w\right) \cdot C_t + m_g \cdot F_n \cdot H}{F_n \cdot \left(H + \frac{I_2}{m \cdot r_{e,0}} + \frac{F_n}{K_x}\right)}; \quad S_t = 0. \tag{18}$$

The calculations with any combination of initial conditions, can be done with the use of the appropriate surface. The system of differential equations can be integrated with the aid of the Runge-Kutta method. The results acquired from the mathematical model of two-free-axis are similar to those which were obtained when the flexibility of the abrasive layer was not taken into account.

6. Conclusions

The adoption of the flexibility in the model resulted in the appearance of the angular parameters of the oscillations of the wheel during its full turn-over. The oscillation frequency of the mathematical model was the same as that of the experimental device, which proves the existence of motion between the wheel-base and the tyre and at the same time presumes the flexibility of the tyre.

The effect of the above mentioned flexibility of the tyre on the dynamics of the rolling of the braked wheel may decrease the efficiency of the ABS and braking system of the aircraft. The dynamical influences related to tyre flexibility can be made perceptible during the design of such systems only with the aid of the two-free-axis model [4], [5].

References

[1] BREWER, H. K.: Parameters Affecting Tyre Control Forces, IAII Pap. 1974, No. 966, pp. 1–9.

[2] HAINLINE, B. C. – AMBERG, R. L. – SRINATH, S. K.: Prediction of Aircraft Braking Friction on Wet Runways, A Look at Past and Current Research Activities, *SAE Tech. Pap.* Ser. 1983, No. 831562, pp. 1–15.

[3] KISS, L.: Lateral Motion of Aircraft on the Runway, *19th Congress of the International Council of the Aeronautical Sciences*, pp. 1364–1371; 18–23, September, 1994, Aneheim, California, USA.

[4] KISS, L.: Some Questions about Operation of Aircraft on Ground Airfield, Doct. Univers. Thesis, Military Academy, Oct. 1994, Budapest, Hungary.

[5] WATTLING, A. G.: The Dynamic Response of an Aircraft Wheel to Variations in Runway Friction, PH.D. Thesis, Department of Aerospace Engineering, University of Bristol, March, 1988.

PERIODICA POLYTECHNICA SER. TRANSP. ENG. VOL. 24, NO. 2, PP. 117-138 (1998)

OPTIMAL CONTROL – OPTIMAL IDENTIFICATION (NEW PARADIGMS – NEW SOLUTIONS)

László KEVICZKY

Computer and Automation Institute
Hungarian Academy of Sciences
H–1111 Budapest, Kende u. 13–17
Phone: +361-1665-435, Fax: +361-1667-503
e-mail: h10kev@huella.bitnet

Received: March 30, 1995

Abstract

A new *generic optimal controller* structure and regulator design method are introduced avoiding the solution of polynomial equations. The model sensitivity properties of some combined identification and control schemes are investigated. It is shown that a new structure is superior to the others. An applicable strategy for iterative control refinement based on the *generic scheme* is presented and illustrated by simulation examples. A worst-case optimal input design algorithm is also introduced to increase the robustness of the closed-loop control in the relevant medium frequency range by generating a 'maximum-variance' reference signal. The adaptive version of the control refinement strategy is also shown with a special 'triple-control' extension for recursive optimal input design.

1. Introduction

The need to design high performance control systems has not lost the importance inspite of the thousands of methods and algorithms published in the past decades. The huge number of papers indicates that no unique or best method was found. The solution depends on the model, criterion, uncertainties, disturbances, constraints, etc. (sometimes even on the designer's taste).

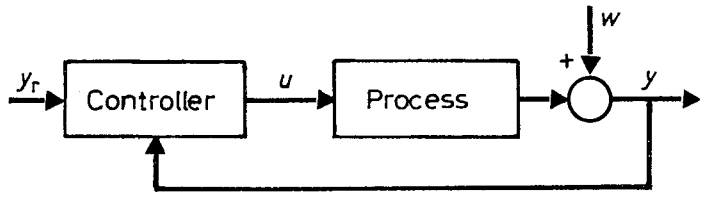


Fig. 1. A general closed-loop control system

A general closed-loop system is shown in *Fig. 1,* where y_r, u, y and w are the reference, input, output and disturbance signals, respectively. Here discrete-time representations are considered for computer controlled systems.

The argument k of variables means the integer value discrete time (integer multiple of the sampling period) and z^{-1} means the backward shift operator $(z^{-1}y(k) = y(k-1))$.

Many experts believe that the essence of all control problems can be led back for the simple problem shown in *Fig. 2*, i.e., how to choose a serial compensator transfer function X to S ensuring a unity dynamic transfer. The trivial solution $X = S^{-1}$ is not always applicable because S is not invertible. This is mostly the case if the control is discrete time and based on sampled linear dynamic systems. In a general case the system $S = S_+S_-$ is factorable for inverse stable S_+ and inverse unstable S_- components. Because S_- cannot be eliminated by simple cancellation mechanism it is sometimes called invariant system component. A heuristic but widely applicable solution is to choose $X = S_+^{-1}$, when the inverse stable part is cancelled, however, the invariant inverse unstable factor is untouched. Several controller design schemes are based on this method, inspite of there is no optimality connected. Unfortunately, the remaining invariant factor can sometimes cause not tolerable dynamics, so its effect must be attenuated. This can be done, if we use a criterion for this purpose. It is known from the classical Wiener framework of optimal stochastic systems, that the solution of the minimum mean square error (H_2) problem

$$X = \arg\min_X \left\{E\left\{\varepsilon^2(k)\right\}\right\} = \arg\min_X \left\{E\left\{\left[(1 - XS_+S_-)\, n(k)\right]^2\right\}\right\} \qquad (1)$$

can be obtained if

$$X = \widetilde{S}_-^{-1} S_+^{-1}, \qquad (2)$$

where $\widetilde{S}_-$ is obtained by reflecting the zeros (they are unstable!) of S_- through the unit circle and providing

$$\widetilde{S}_-^{-1}(1) S_-(1) = 1. \qquad (3)$$

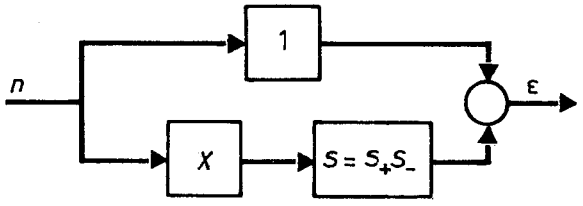


Fig. 2. The simple problem of control systems

Note that the solution depends on the applied input excitation $n(k)$. Here $n(k)$ is assumed as a white noise sequence.

A characteristic approach of optimal controller schemes is called *pole-placement technique* (ÅSTRÖM – WITTENMARK (1984); LANDAU (1990))

targeting to provide prescribed transient properties for the servo and disturbance rejection paradigm of closed-loop controller design. The standard technique representing a two degrees of freedom so-called $\mathcal{R} - \mathcal{S} - \mathcal{T}$ controller assumes basically the structure shown in *Fig. 3.* Here $\mathcal{R}$, $\mathcal{S}$ and $\mathcal{T}$ are polynomials. The advantage of this scheme is that the implementation of a so-called direct adaptive regulator method is very easy, because it is easy to construct a predictor equation linear in the parameters of these polynomials. The disadvantage of this scheme is that it hides the internal operation of an optimal system and special considerations are necessary in a recursive parameter estimation algorithm because the $\mathcal{R}$, $\mathcal{S}$ and $\mathcal{T}$ are redundant, having more parameters than minimally necessary, furthermore the solution of a Diophantine equation is necessary to obtain the regulator polynomials.

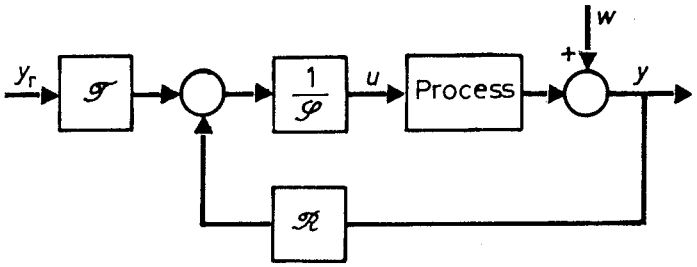


Fig. 3. A classical pole-placement controller

Another well-known classical scheme of optimal control systems is called *internal model principle.* The name originates from the system model applied in the controller. This scheme has a much less known form if we want to use the same principle for inverse unstable factors. This *modified internal model principle* is shown in *Fig. 4,* if only inverse stable factors are cancelled. Note that the whole system should be taken into consideration in the internal model, because realizability problem arises only using the inverted model. Here P_r and P_w are the desired overall tracking and disturbance rejection transfer functions (or reference models) for the design requirements. They can also be interpreted as predictors for the reference and output disturbance signals. In this case, e.g., $\widehat{w}$ is the estimated (or predicted) disturbance. (An ideal case was assumed above when the true process S is known to ease the understanding of the basic schemes. This assumption is good to explain the operation of the system, however, only a process model M is available in most practical cases as it will be discussed later.) If we want to attenuate the invariant system component then $\widetilde{S}_-^{-1} S_+^{-1}$ must be used instead of S_+^{-1} in the *partially inverse model,* according to the optimality of *Eq.* (2).

The advantage of this optimal control scheme is that the principal operation of the regulator is very easy to follow and the computations of the regulator polynomials are easy and obvious. The disadvantage of this

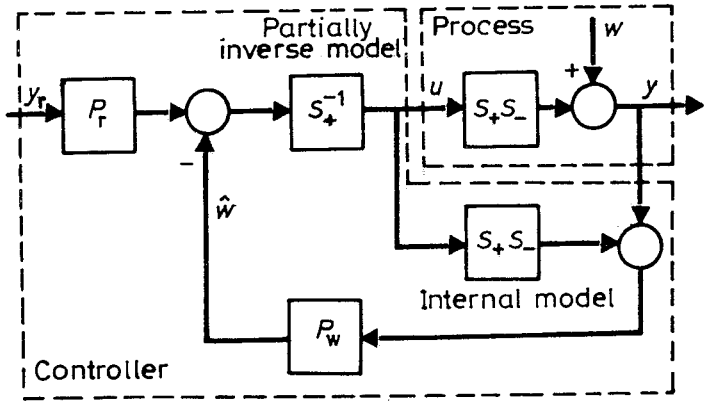


Fig. 4. Modified internal model principle for inverse unstable factor

scheme is that the identification (parameter estimation) method indicated by the *internal model* in the closed loop arises several difficulties.

Both above approaches have the general problems of the known optimal controller structures that the *identification and control errors are different.* Therefore these schemes are not the best ones for developing and analyzing simultaneous identification and (adaptive) control algorithms.

In this paper a new *generic structure* (KEVICZKY – BÁNYÁSZ (1994)) is discussed which allows a very simple procedure to design optimal control systems when the identification and control errors are identical. Their relationships to the previously discussed schemes are also presented.

2. A New Controller Structure

Let us introduce another new structure shown in *Fig. 5* to design optimal controllers. Here S is the system, R is the regulator, $\tilde{P}_r$ is a precompensator transfer function. The system output

$$y(k) = \tilde{P}_r S y_r(k) + \frac{1}{1+RS} w(k) = y_t(k) + y_d(k) \tag{4}$$

is very special, because in spite of the closed-loop the tracking behaviour $\left(y_t(k) = \tilde{P}_r S y_r(k)\right)$ is independent of the regulator R. This structure practically opens the loop for the command signal and the selected feedforward compensator (or observer)

$$\tilde{P}_r = P_r S^{-1} \tag{5}$$

provides a desired

$$y_t(k) = P_r y_r(k) \tag{6}$$

tracking (servo) transfer function by P_r. Observe that selecting a regulator

$$R = \frac{P_w}{1 - P_w} S^{-1} = C S^{-1} \tag{7}$$

a desired regulating (or disturbance rejection) behavior

$$y_d(k) = (1 - P_w)\, w(k) \tag{8}$$

can be reached by P_w. The P_r and P_w transfer functions contain the desired poles to be placed, so they can be called as reference models. Note that arbitrary zeros can also be placed, however, the calculation of the precompensator $\tilde{P}_r$ and regulator R requires the inverse of the process S. (Here also the ideal case was assumed when the true process S is known. If P_w is the best LS predictor of w, then the obtained regulator is the minimum variance (MV) regulator.) It is easy to prove that this scheme can be rearranged to the scheme in *Fig. 2* by straightforward block manipulations. Assuming that

$$S = \frac{B}{A}, \tag{9}$$

$$P_r = \frac{B_r}{A_r} \qquad \text{and} \qquad P_w = \frac{B_w}{A_w} \tag{10}$$

the classical pole-placement regulator polynomials can also be easily computed according to the previous formulas and they are given by

$$\mathcal{T} = A A_w B_r, \tag{11}$$

$$\mathcal{S} = A_r B\,(A_w - B_w)\,, \tag{12}$$

$$\mathcal{R} = A A_r B_w. \tag{13}$$

(Observe the common factors that make the parametrization for a direct self-tuning technique redundant and the recursive estimation difficult.

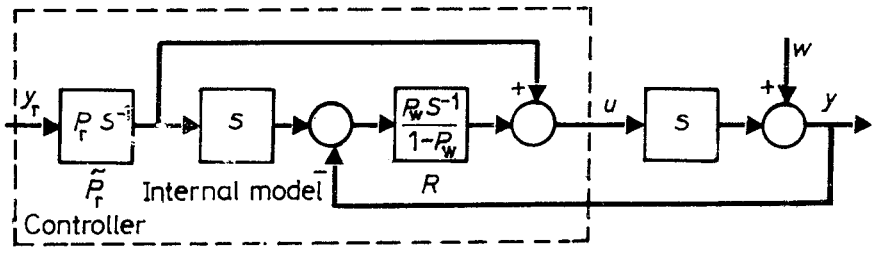


Fig. 5. The new control system structure

If we use the above precompensator and regulator the controller shown in *Fig. 5* can be further simplified according to *Fig. 6.* The advantage of this

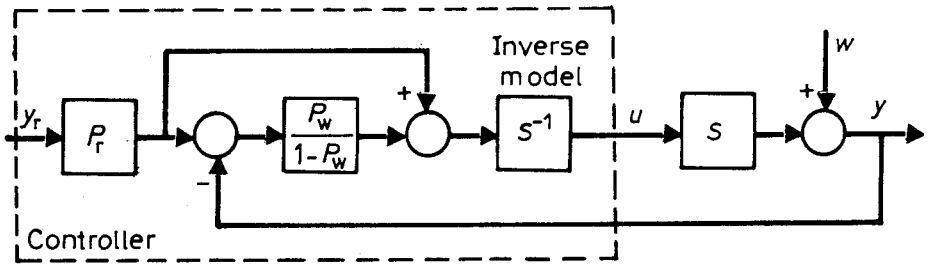


Fig. 6. Equivalent form of the new controller

approach is that the three transfer functions of the controller can directly be computed without solving the usual polynomial equation. A further and very important advantage is that the identification and control errors are identical, so this scheme is especially good for combined identification and control problems. (The very limited usability for inverse stable systems will be eliminated in the next section.) It is easy to show that *Fig. 5* can be rearranged to the schemes either in *Fig. 3* or in *Fig. 4* by straightforward block manipulations. So the *classical pole-placement regulator scheme and the internal model principle scheme are identical* to each other and to the new scheme if the appropriate transfer functions are selected.

3. A Generic Scheme for Optimal Pole-Placement Controllers

Because the above method is based on full pole and zero cancellation the extended applicability can be reached by partial cancellation of those system (model) components only, which are inverse stable. Assume that

$$S = S_+ S_- = S_+ \overline{S}_- z^{-d}, \tag{14}$$

where S_+ means the inverse stable factor. Here $S_- = \overline{S}_- z^{-d}$ is the non-invertible part where the discrete time delay z^{-d} is also a factor whose inverse z^d is not realizable. In case of a partial cancellation we should use the precompensator

$$\tilde{P}_r = P_r S_+^{-1} \tag{15}$$

instead of (5). This results in

$$y_t(k) = P_r S_+^{-1} S_+ \overline{S}_- z^{-d} y_r k = P_r \overline{S}_- z^{-d} y_r k. \tag{16}$$

Selecting the regulator as

$$R = \frac{P_w S_+^{-1}}{1 - P_w \overline{S}_- z^{-d}} = C S_+^{-1} \tag{17}$$

one can obtain the regulatory transfer as

$$y_d(k) = \left(1 - P_w \overline{S}_- z^{-d}\right) w(k). \tag{18}$$

Eqs (16) and (18) show that in case of partial cancellation, which is the general case, when we have inverse unstable system factors (e.g. nonminimum phase systems) or time delay we cannot reach the ideal servo (P_r) and disturbance rejection ($1 - P_w$) transfer functions only their modified $\left(P_r \overline{S}_- z^{-d}\right)$ and $\left(1 - P_w \overline{S}_- z^{-d}\right)$ forms. Note that the modifications do not depend on us, instead they depend on the system itself only. Therefore $\overline{S}_-$ is sometimes called invariant process factor (mostly zeros). (Because reasonably $P_r(1) = 1$ and $P_w(1) = 1$ are selected, it is also reasonable to choose $\overline{S}_-(1) = 1$ in the factorization *Eq.* (14)). *Fig. 7* shows the transfer functions in the new controller scheme, which can even be called a *generic* optimal pole-placement controller in this case. This scheme can be further simplified according to *Fig. 8.* It is easy to show that *Fig. 7* can be rearranged to the schemes either in *Fig. 3* or in *Fig. 4* by straightforward block manipulations. So the *classical pole-placement regulator scheme* and the *modified internal model principle scheme* are identical to each other and to the *new generic scheme* if the appropriate transfer functions are selected.

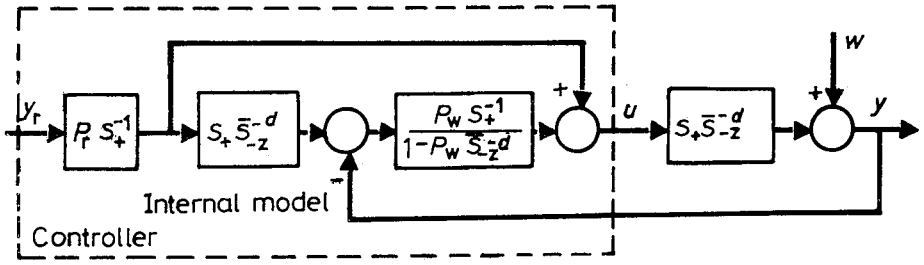


Fig. 7. The *generic pole-placement controller*

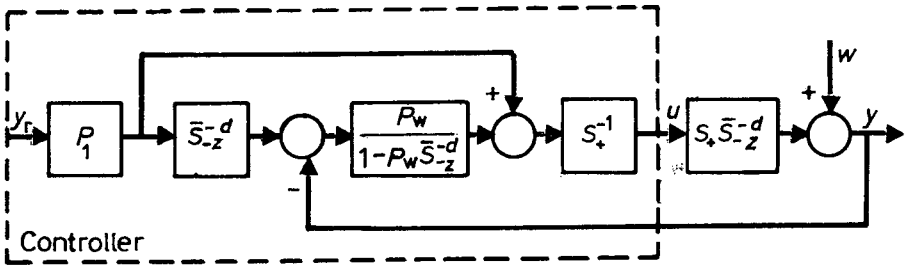


Fig. 8. The equivalent form of the *generic controller*

Assuming that instead of (9) the system transfer function corresponding to (14) is given by

$$S = \frac{B_+ B_-}{A} z^{-d} \tag{19}$$

only the polynomial $\mathcal{S}$ will be changed

$$\mathcal{S} = A_r B_+ \left(A_w - B_w B_- z^{-d} \right) \tag{20}$$

and the computation of $\mathcal{T}$ and $\mathcal{R}$ remains according to (11) and (13).

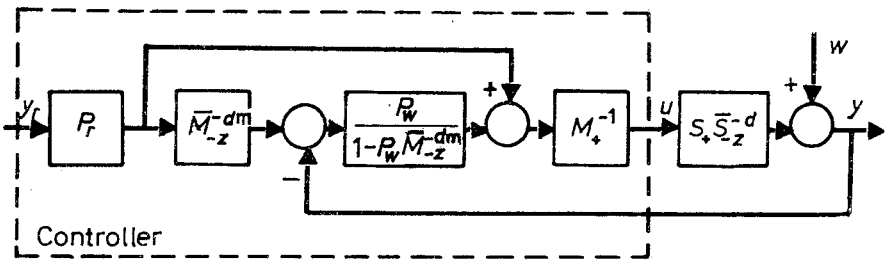


Fig. 9. The *generic* model based pole-placement controller

The final optimal controller is already very general because it covers the most critical processes where the design is not trivial. In the practice we should rather use the model M instead of the system S in the internal model, so the design procedure should apply

$$M = M_+ \overline{M}_- z^{-d_m} = \left(\frac{\widehat{B}_+}{\widehat{A}} \right) \left(\widehat{B}_- \right) z^{-d_m} = \frac{\widehat{B}_+ \widehat{B}_- z^{-d_m}}{\widehat{A}} = \frac{\widehat{B}}{\widehat{A}} z^{-d_m} \tag{21}$$

instead of (14). Here M_+, $\overline{M}_-$ and z^{-d_m} are the inverse stable, inverse unstable factors and the delay time of the model, respectively. So the optimal controller will change according to *Fig. 9.* This controller is very easy to implement in a computer controlled system and it keeps the advantage of the original idea of the new controller structure, i.e., it does not require the solution of a polynomial (Diophantine) equation to obtain the controller transfer functions or polynomials, instead (11), (13) and (20) should be applied now, where $\widehat{B}_+ \widehat{B}_- z^{-d}$ and $\widehat{A}$ are the numerator and denominator polynomials of the model M instead of the system S, as *Fig. 10* shows.

It was already mentioned that the not cancellable factors of the system will be factors of the modified reference models, so they are invariant to any control strategy. However, their influence on the original reference models can be minimized, if we use a criterion for this purpose. Using the Wiener design principle shown in the introduction the precompensator

$$\widetilde{P}_r = P_r \widetilde{\overline{S}}_-^{-1} S_+^{-1} \tag{22}$$

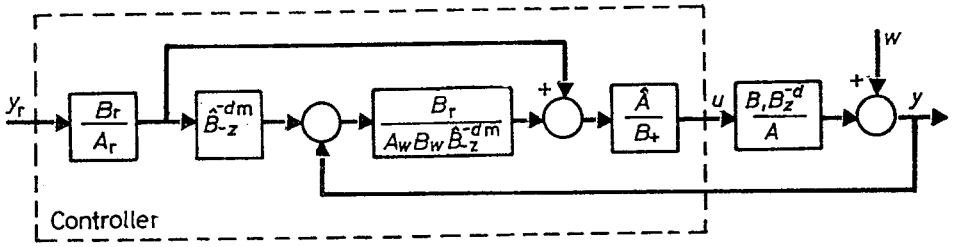


Fig. 10. The polynomial design of the *generic* pole-placement controller

and the regulator

$$R = \frac{P_w \tilde{\overline{S}}_-^{-1} S_+^{-1}}{1 - P_w \tilde{\overline{S}}_-^{-1} \overline{S}_- z^{-d}} = C S_+^{-1} \tag{23}$$

must be used and the overall system equation

$$y(k) = P_r \tilde{\overline{S}}_-^{-1} \overline{S}_- z^{-d} y_r(k) + \left(1 - P_w \tilde{\overline{S}}_-^{-1} \overline{S}_- z^{-d}\right) w(k) \tag{24}$$

is obtained. *Fig. 11* shows the practical realization of this strategy if the model of the system is used.

Note that the application of the compensator $\tilde{\overline{S}}_-^{-1}$ on $\overline{S}_-$ is optimal for a given excitation only, in this case for white noise (or approximately for wide bandwidth) disturbance. Therefore, the optimal pole-placement controller shown in *Fig. 8* is suggested for most practical applications.

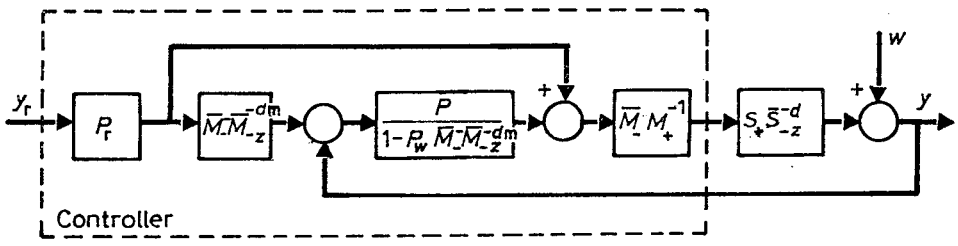


Fig. 11. Attenuating the invariant factors by Wiener design

4. Combined Identification and Control Schemes

Since the beginning the key paradigm of designing control systems is how to handle uncertainties associated with the plant. One of the main techniques is adaptive control intending to learn parameter and disturbance uncertainties

in varying circumstances. Another important approach is to implement robustness features at simultaneous identification and control procedures.

In the traditional approach to analysis and design of an adaptive control system the unknown plant is represented by a model, which is almost known except a few parameters assumed time varying. Having the estimated parameters the controller is updated according to the certainty equivalence principle. The unstructured uncertainties are mostly ignored in these cases, therefore these adaptive regulators are not robust. To tell the truth it is not easy at all to consider these uncertainties at the classical adaptive systems and to guarantee proper transients during the learning adaptation phase or abrupt parameter changes.

In a classical robust control approach the regulator is designed on the basis of a nominal model of the plant associated with the associated parametric and unstructured model uncertainties explicitly taking into account. (Unfortunately, mostly analytical forms are required which are very rarely available at practical applications, except a few special real cases or the examples 'god given' apriori information). Here stability robustness is guaranteed and performance robustness is achieved sometimes. The weakness of this approach is that it considers only the apriori information on the model and neglects that the characteristics of the plant could be learnt even when it is controlled. Therefore classical robust control approaches mostly result in a conservative design in terms of performance.

(In Section 8 a new approach will be introduced, when the uncertainty of the model coming from the parameter estimation will be minimized in the relevant medium frequency range around the cross-over frequency.)

4.1. Open-loop Identification and Closed-loop Control

The simplest strategy to combine modelling and control if the identification is performed in an open-loop experiment to obtain an optimal model M^*, selected from a model class $\mathcal{M}$, by minimizing an identification criterion $Q_{i0}(\varepsilon_0)$ function of the open-loop identification error ε_0, i.e.,

$$M^* = \arg \min_{M \in \mathcal{M}} Q_{i0}(\varepsilon_0) = \arg \min_{M \in \mathcal{M}} Q_{i0}(M, \mathbf{x}, \Delta) = M_0^*(\mathcal{M}, \mathbf{x}, S), \quad (25)$$

where $\mathbf{x} = \{x(k); k = 1, \ldots N\}$ is the applied input excitation series and

$$\Delta = S - M \quad (26)$$

is the additive model uncertainty between the system S and model M (*Fig. 12*). The optimal regulator R^*, selected from a regulator class $\mathcal{R}$, is obtained by minimizing a control criterion $Q_c(e_c)$ function of the closed-loop control error e_c (*Fig. 12*), i.e.,

$$R = \arg \min_{R \in \mathcal{R}} Q_c(e_c) = \arg \min_{R \in \mathcal{R}} Q_c(R, S) = R^*(\mathcal{R}, S). \quad (27)$$

(Note that the open-loop input excitation x is different from the closed-loop input u acting on the process.) Because only the model of the system is known in a practical case, therefore the most widely applicable strategy is to substitute S by M in (27). This strategy is called the *certainty equivalence principle* and realized by

$$R = \arg\min_{R\in\mathcal{R}} \widehat{Q}_c\,(\widehat{e}_c) = \arg\min_{R\in\mathcal{R}} \widehat{Q}_c(R, M) = \widehat{R}^*(\mathcal{R}, M). \tag{28}$$

In this case the optimal regulator does not reach the theoretical optimum, because the model M is used instead of the system S. However, it is possible to form an iterative control refinement procedure improving the model and regulator step by step:

1. Identify the model using the modelling step

$$M_i = M_0^*\,(\mathcal{M}, \mathbf{x}_{i-1}, S)\,. \tag{29}$$

2. Calculate the optimal regulator

$$R_i = \widehat{R}^*\,(\mathcal{R}, M_i)\,. \tag{30}$$

3. Determine an optimal input excitation for the open-loop identification

$$\mathbf{x}_i = \widehat{\mathbf{x}}_0^*\,(\mathcal{H}, M_i, R_i) = \mathcal{D}_{x\in\mathcal{H}}\,(\mathcal{H}, \mathbf{x}, M_i, R_i)\,. \tag{31}$$

   Here $\mathcal{H}$ is the (mostly amplitude or energy) constrained input signal domain. This step is sometimes called optimal input design and the operation is denoted by $\mathcal{D}\,(\ldots)$.

4. Once M_i and R_i are found we can continue to increase the closed-loop bandwidth repeating the procedure. The iterative process is continued from step 1, while a stop condition is not fulfilled (until the ultimate control objective is achieved or it is terminated because of reaching some vital constraints).

For comparison it is interesting to see that in the above case the identification and control errors are

$$\varepsilon_0 = (S - M)x = m\left(\frac{\Delta}{M}\right)x = H_0\left(\frac{\Delta}{M}\right)x \tag{32}$$

and

$$e_c = \frac{1}{1+RS}\widehat{y}_r + \frac{1}{1+RS}w; \qquad \widehat{y}_r = P_r y_r, \tag{33}$$

where P_r is the reference model, $\widehat{y}_r = P_r y_r$ is the desired process output (model output or predicted reference signal) and w is the output disturbance.

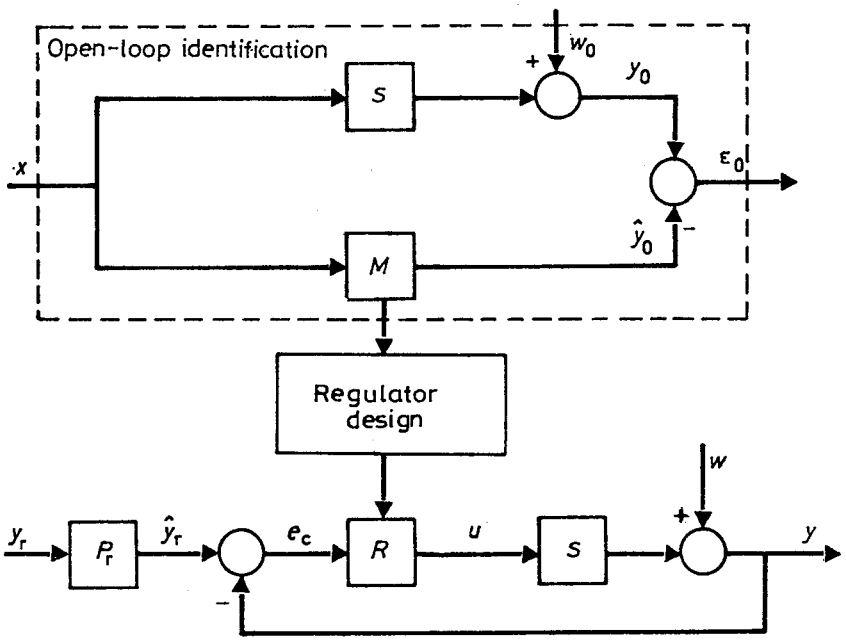


Fig. 12. Combination of *open-loop identification* and *closed-loop control*

4.2. Closed-loop Identification and Closed-loop Control

The formal description of the procedure how to combine modelling and control if the identification is performed in a closed-loop experiment is very similar to the previous section, however, the identification criterion $Q_{ic}(\varepsilon_c)$ is now the function of the closed-loop identification error ε_c, i.e.

$$M^* = \arg\min_{M\in\mathcal{M}} Q_{ic}(\varepsilon_c) = \arg\min_{M\in\mathcal{M}} Q_{ic}(M, R, y_r, \Delta) = M_c^*(\mathcal{M}, y_r, S), \tag{34}$$

and the regulator is designed again by *Eq.* (28). It is possible to form different combined schemes depending on the structure of the optimal control and the combination of the sequential identification and control steps. In this sequential procedure it is a general observation that

1. The human first learns to control over a limited bandwidth, and learning pushes out the bandwidth over which an accurate model of the plant is known.
2. The human first implements a low gain controller, and learning allows the loop to be tightened.

On the basis of these observations an adaptive robust control philosophy, the *windsurfer approach,* was proposed by ANDERSON – KOSUT (1991). They

use a *parallel closed-loop optimal controller scheme,* which is very widely used in the analysis of control relevant identification and iterative control refinement procedures. Note that the classical so-called direct adaptive control algorithms generally use a somewhat different scheme which is called as *parallel in-loop optimal controller scheme.* We will analyze these schemes and the new *generic optimal controller scheme* also in the sequel. For the sake of simplicity this comparison will be discussed here for inverse stable processes first.

The generic optimal controller scheme. The *generic optimal controller scheme* (KEVICZKY – BÁNYÁSZ (1994)) was shown in Section 2.

The structure of the optimal controller gives a special insight to understand the operation of a feedback loop for the servo and disturbance rejection paradigm. It is easy to see the role of the knowledge of the model of the system and the role and appearance of the factors of the system that are invariant to any control strategy. A little bit modified form of the *generic scheme* of *Fig. 5* will be used here as *Fig. 13* shows.

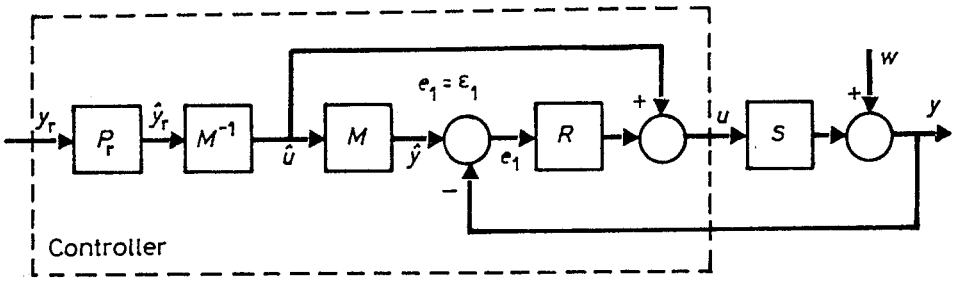


Fig. 13. The *generic optimal controller scheme*

This structure has further advantages in handling system uncertainties and a new canonic sensitivity scheme can also be obtained. This *generic scheme* opens a new way to analyze combined modelling and control issues. *Fig. 13* is the long searched ideal scheme for the combined identification and control problem, because in this case *the control and modelling errors are identical.* So the new scheme provides an excellent possibility also to study robust identification for control.

The common identification and control errors are

$$e_1 = \varepsilon_1 = \frac{1}{1+RS}\left(\frac{\Delta}{M}\right)\hat{y}_r - \frac{1}{1+RS}w = -\left[H_1\left(\frac{\Delta}{M}\right)\hat{y}_r + E_1 w\right]. \quad (35)$$

5. On the Generic Optimal Controller Scheme

Inverse stable processes. It was presented above that the *generic optimal controller scheme* has certain advantages comparing to others shown. It

is also interesting to show that the common modelling and control error $e_1 = \varepsilon_1$ given by (35) can be also expressed as

$$e_1 = \varepsilon_1 = -\left[H_1\left(\frac{\Delta}{M}\right)\hat{y}_r + E_1 w\right] = -[(Su + w) - M\hat{u}] = -(y - \hat{y}). \quad (36)$$

This form and *Fig. 13* shows an obvious way how to perform the identification step in a combined identification and control scheme, i.e., we should use a regular identification algorithm based on the auxiliary variable $\hat{u}$ and the measured controlled variable y as *Fig. 14* shows. Note that $\hat{u}$ and y must be obtained from the closed-loop operated by the generic optimal controller structure.

Because $\hat{u}$ depends on the model M only an iterative control refinement procedure can be performed. Its simplest – so-called relaxation type – iteration can be formed in the following way for an off-line case using N samples (i-th iteration is shown):

1. Calculate the auxiliary variable $\hat{u}_i$ based on the available model M_{i-1}, the reference model P_r and the applied reference signal series $y_r^i = \{y_r^i(k);\ k = 1, \ldots, N\}$

$$\hat{u}_i(k) = M_{i-1}^{-1} P_r y_r(k); \qquad k = 1, \ldots, N. \quad (37)$$

   Here i denotes the index of the iteration and note that y_r^i does not necessarily change by iteration.

2. Identify a model between $\hat{\mathbf{u}}_i$ and $\mathbf{y}_i$ using the modelling step

$$\begin{aligned} M_i &= \arg\min_{M \in \mathcal{M}} Q_{ic}(\mathcal{M}, \Delta, \hat{\mathbf{u}}_i, R_{i-1}, P_w) = \\ &= M_c^*(\mathcal{M}, M_{i-1}, R_{i-1}, \mathbf{y}_r, S, P_r, P_w) \end{aligned} \quad (38)$$

$$(\hat{\mathbf{u}}_i = \{\hat{u}_i(k);\ k = 1, \ldots, N\}; \qquad \mathbf{y}_i = \{y_i(k);\ k = 1, \ldots, N\}).$$

3. Calculate the optimal regulator based on (7)

$$R_i = CM_i^{-1} = \frac{P_w}{1 - P_w} M_i^{-1} \quad (39)$$

   the compute the process input u_i as

$$u_i(k) = P_r\left(R_i + M_i^{-1}\right) y_r^i(k) - R_i y_i(k); \qquad k = 1, \ldots, N \quad (40)$$

   and apply to the process.

4. (There is a possibility to optimize the applied reference signal series in this step by a proper input design procedure.)

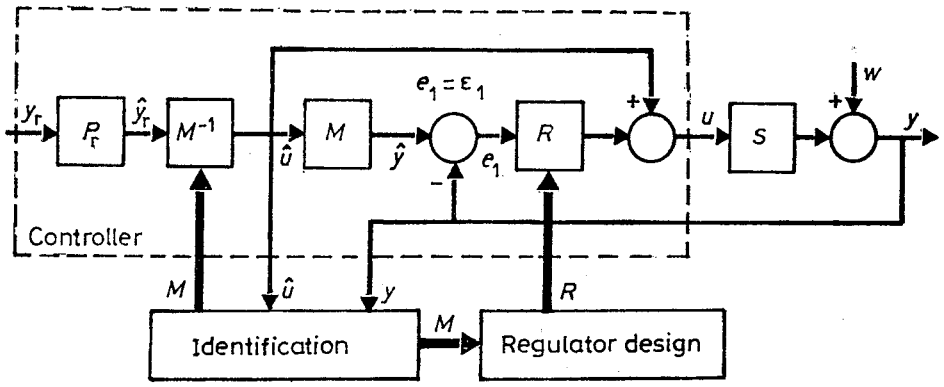


Fig. 14. Identification and regulator design at the *generic optimal controller scheme*

5. Once M_i and R_i are found we can continue to increase the closed-loop bandwidth repeating the procedure. The iterative process is continued from step 1, while a stop condition is not fulfilled (until the ultimate control objective is achieved or it is terminated because of reaching some vital constraints).

Inverse unstable processes. Note that both $\hat{u}$ and R need the inverse of M. Since this method is based on full pole and zero cancellation the extended applicability can be reached by partial cancellation of those system (model) components only, which are inverse stable. This *extended generic optimal controller scheme* is shown in *Fig. 15* which corresponds to *Fig. 7.* The identification and control errors are identical at this scheme, too, i.e.,

$$\begin{aligned} e_1 &= \frac{-\Delta P_r M_-}{(1+RS)M_+M_-} y_r - \frac{1}{1+RS} w = \\ &= \frac{-1}{1+RS}\left(\frac{\Delta}{M_+}\right)\hat{y}_r - \frac{1}{1+RS} w = \\ &= -\left[H_1\left(\frac{\Delta}{M_+}\right)\hat{y}_r + E_1 w\right] = \varepsilon_1. \end{aligned} \tag{41}$$

The off-line iterative control refinement procedure described by *Eqs* (37), (38), (39) and (40) and steps 1 – 5 can also be applied here, if instead of these equations the corresponding

$$\hat{u}_i(k) = \left(M_+^{i-1}\right)^{-1} P_r y_r(k); \qquad k = 1, \ldots, N, \tag{42}$$

$$R_i = \frac{P_w}{1-P_w}\left(M_+^i\right)^{-1} = C\left(M_+^i\right)^{-1}, \tag{43}$$

$$u_i(k) = P_r \left(R_i M_-^i + \left(M_+^i \right)^{-1} \right) y_r^i(k) - R_i y_i(k); \qquad k = 1, \ldots, N \quad (44)$$

formulas are used.

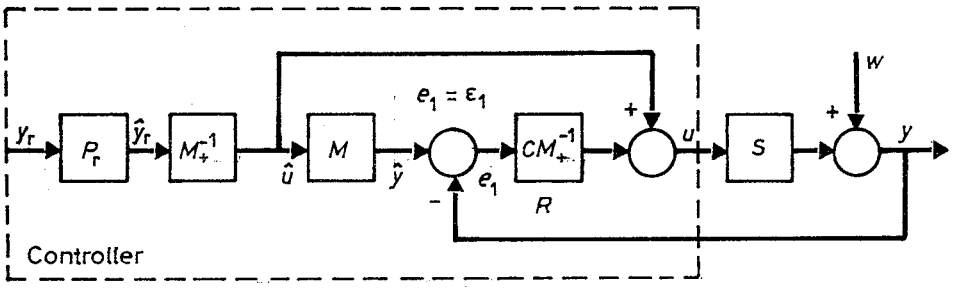


Fig. 15. The *extended generic optimal controller scheme* for inverse unstable processes

6. Examples for Off-line Iterative Regulator Refinement

Example 6.1 Let the process be given by

$$S = \frac{0.007869}{1 - 0.606531 z^{-1}} z^{-4} \quad (45)$$

which is a sampled-time (sampling time is $h = 0.05$ s and $d = 4$) first order approximation of a helicopter 'stick-input/roll-rate-output' model. Apply the unity gain tracking and disturbance rejection reference models

$$P_r = \frac{0.5 z^{-1}}{1 - 0.5 z^{-1}} \qquad \text{and} \qquad P_w = \frac{0.2 z^{-1}}{1 - 0.8 z^{-1}}, \quad (46)$$

and start the iterative control refinement by the model

$$M_0 = \frac{0.01}{1 - 0.4 z^{-1}} z^{-4}, \quad (47)$$

i.e., $d_m = d = 4$. *Fig. 16* shows the control and identification loss functions (variances) by the iteration. It can be well seen that the iteration is quite fast reaching the optimal values after 4 steps.

A unity amplitude square wave reference input signal with periodic time 40 samples was applied and the off-line procedure used $N = 100$ samples. In the simulation runs an additive white noise was used as output disturbance with a standard deviation $\lambda = 0.01$. We used a simple off-line *LS* method for parameter estimation only to demonstrate the operation of

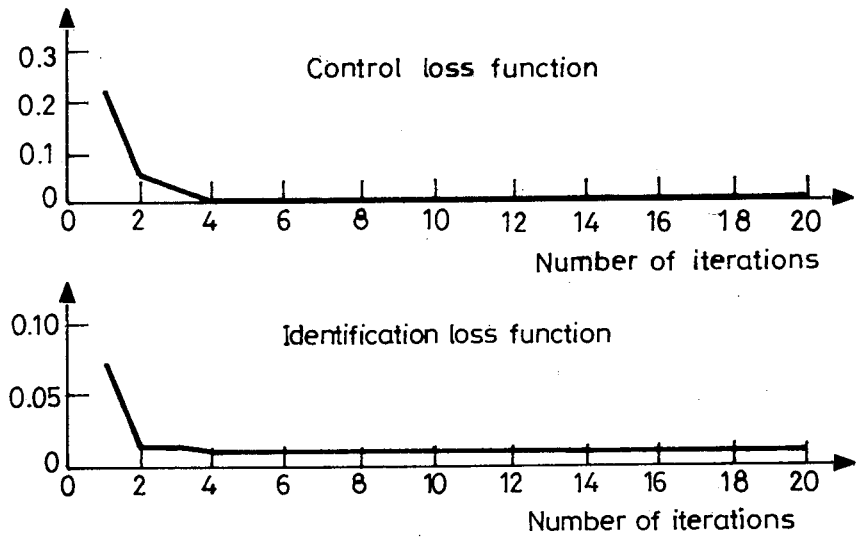


Fig. 16. Loss functions in an iterative control refinement procedure (first order example)

the iterative algorithm. (Our experience showed that this method works very fine in this scheme while the noise level is low. However, in a noisy case the proper parameter estimation method (ELS, ML, etc.) should be applied corresponding to the given process and noise model structure.) The outputs of the tracking reference model (continuous) and the controlled process (dashed) are shown in *Fig. 17* before and after the iterative control refinement.

Example 6.2 Let the process (GEVERS (1991)) be given by

$$S = \frac{0.0364z^{-1}\left(1 + 1.2z^{-1}\right)}{1 - 1.6z^{-1} + 0.68z^{-2}}, \tag{48}$$

where the same reference input and output noise was used as in the previous example with the initial model

$$M_0 = \frac{0.04z^{-1}\left(1 + 1.0z^{-1}\right)}{1 - 1.4z^{-1} + 0.5z^{-2}}. \tag{49}$$

Fig. 18 shows the control and identification loss functions (variances) by the iteration. The outputs of the reference model (continuous) and the controlled process (dashed) are shown in *Fig. 19* before and after the iterative refinement performed by off-line *LS* parameter estimation. Outputs of the controlled process (continuous) and identified model (dashed) before and after the iteration are shown in *Fig. 20.* This figure is a nice example

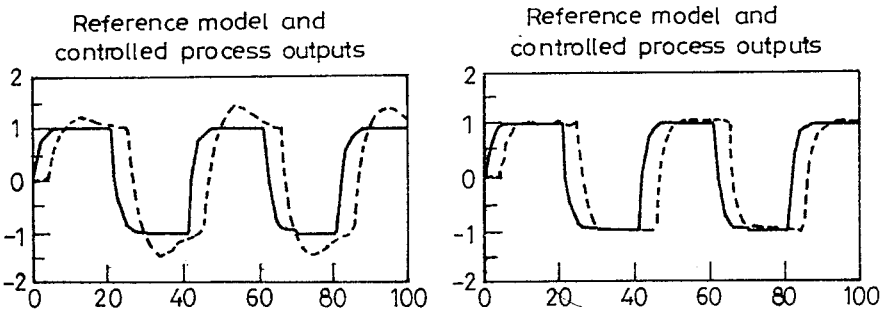


Fig. 17. Outputs of the reference model and the controlled process before and after the iteration (LS)

to explain the necessity for iterative control refinement. One can see that the identified model output fits very nicely to the process output before and after the iteration, too, so the model error is small. However, the control error is very bad before the iteration shown in *Fig. 19.*

7. Adaptive Solution for the On-line Iterative Regulator Refinement

The previous examples demonstrated the nice operation of the off-line iterative control refinement procedure based on the *generic scheme.* Inspite of the good convergence properties the necessary measurements are remarkable. In many applications this is a costly and long procedure to design the optimal regulator. However, it is easy to construct the adaptive control refinement procedure based on the iterative scheme. Following the same steps and properly changing the iteration i to sampling time k the following formulas are obtained (not discussing here the details):

$$\hat{u}(k) = \left(M_+^{k-1}\right)^{-1} P_r y_r(k), \tag{50}$$

$$f\,(k-d_m) = \left[\hat{u}\,(k-d_m)\,, \hat{u}\,(k-d_m-1)\,, \ldots, -y(k-1), -y(k-2), \ldots\right]^T, \tag{51}$$

$$\mathbf{K}_k = \frac{1}{\rho^2}\left\{\mathbf{K}_{k-1} + \frac{\mathbf{K}_{k-1}\mathbf{f}\,(k-d_m)\,\mathbf{f}^T\,(k-d_m)\,\mathbf{K}_{k-1}}{\rho^2 + \mathbf{f}^T\,(k-d_m)\,\mathbf{K}_{k-1}\mathbf{f}\,(k-d_m)}\right\}, \tag{52}$$

$$\hat{\mathbf{p}}_k = \hat{\mathbf{p}}_{k-1} + \mathbf{K}_k\mathbf{f}\,(k-d_m)\left[y(k) - \mathbf{f}^T\,(k-d_m)\,\hat{\mathbf{p}}_{k-1}\right],$$

$$\hat{\mathbf{p}}_k = \left[\hat{b}_0^k,\ \hat{b}_1^k,\ \ldots,\ \hat{a}_1^k,\ \hat{a}_2^k,\ \ldots\right]^T, \tag{53}$$

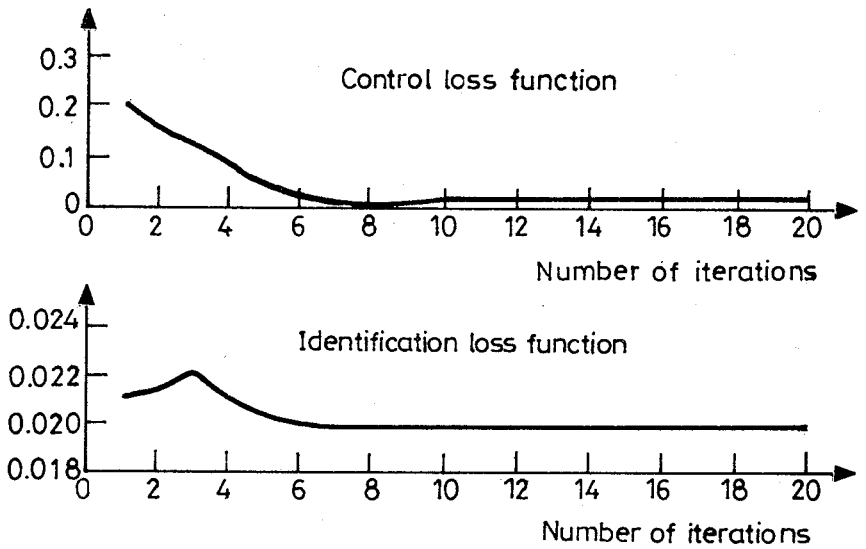


Fig. 18. Loss functions in an iterative control refinement procedure (second order example)

$$R_k = \frac{P_w}{1 - P_w} \left(M_+^k\right)^{-1} = C \left(M_+^k\right)^{-1}, \tag{54}$$

$$u(k) = P_r \left[R_k M_-^k + \left(M_+^k\right)^{-1}\right] y_r(k) - R_k y(k), \tag{55}$$

where ρ is an exponential forgetting factor and *Eqs* (50) – (53) represent a recursive *LS* method (*RLS*) in the simplest so-called naive programming form. Several other recursive parameter estimations can be applied instead of the above *RLS* algorithm.

Note that after having the estimated parameter vector $\hat{\mathbf{p}}_k$ corresponding to (21) obtained M_+ and M_- should be computed by factorization. The above procedure is an on-line strategy performing all refinement steps (50) – (55) in one sampling instance. Here we assumed that the signal $y_r(k)$ itself or its generation rule is given apriori.

8. Adaptive Examples

Example 8.1 Let the process be given by

$$S = \frac{0.001\,(1.1 + 1z^{-1})}{(1 - 1.6693z^{-1} + 0.7788z^{-2})} z^{-2}, \tag{56}$$

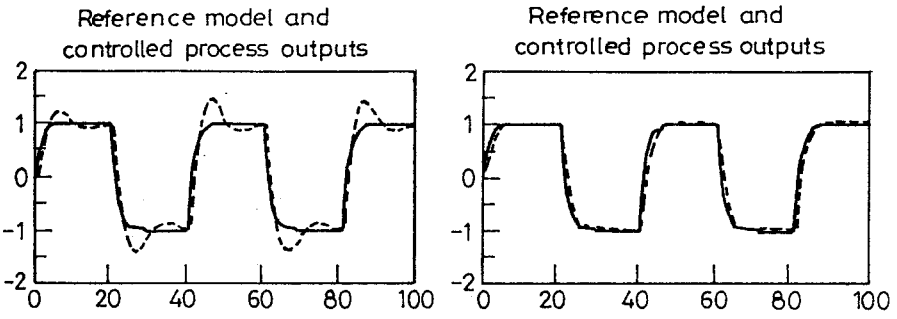


Fig. 19. Outputs of the reference model and the controlled process before and after the iteration (LS)

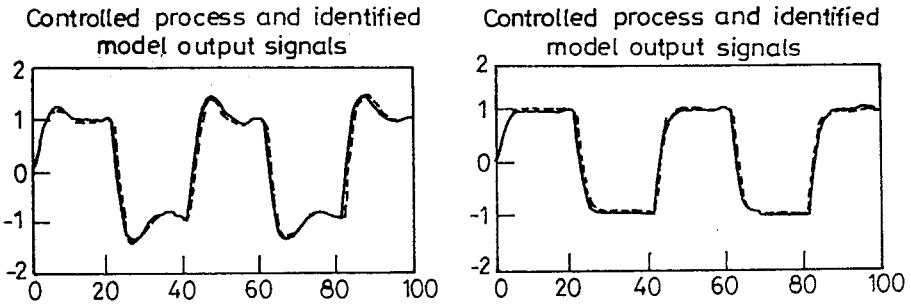


Fig. 20. Outputs of the controlled process and identified model before and after the iteration (LS)

which is a sampled-time (sampling time is $h = 0.05$ s and $d = 2$) second order approximation of a helicopter 'stick input/roll rate output' model. Here the same P_r, P_w, λ and square wave y_r excitation was used as in Example 6.1. The initial model was

$$M_0 = \frac{0.001\,(20 + 0.5z^{-1})}{(1 - 1.5z^{-1} + 0.8z^{-2})} z^{-2}, \tag{57}$$

i.e., $d_m = d = 2$. The outputs of the reference model (continuous) and the controlled process (dashed) furthermore the control (continuous) and modelling (dashed) error signals are shown in *Fig. 21* for $N = 200$ samples using the adaptive control refinement strategy (50) – (55) with $\mathbf{T}_0 = 100\mathbf{I}$ and $\rho^2 = 0.95$.

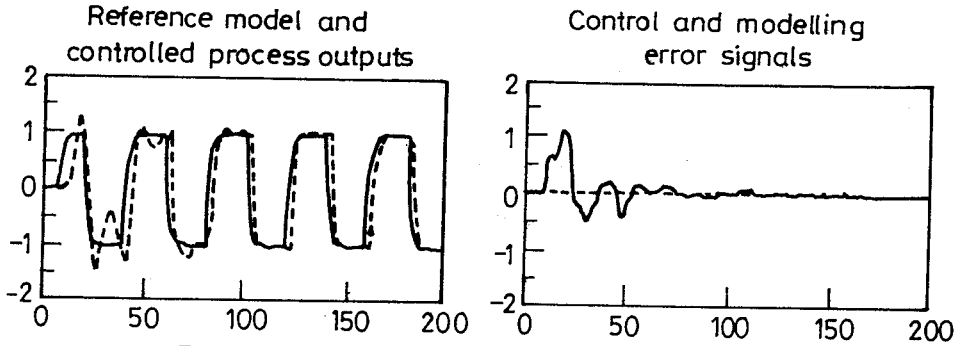


Fig. 21. Operation of the adaptive control refinement

9. Conclusions

The paper introduced a new structure to design optimal pole-placement controllers. This new scheme allows to avoid the explicit solution of a polynomial equation obtaining the transfer function elements of the optimal controller directly. The new design principle is quite general and applicable for nonminimum phase (inverse unstable) and delay time systems, too. The controller is easy to be implemented in computer controlled systems. The structure of the optimal controller gives a special insight to understand the operation of a feedback loop for the servo and disturbance rejection paradigm. It is easy to see the role of the knowledge of the system model and the role and appearance of the system factors that are invariant to any control strategy.

This structure has further advantages in handling system uncertainties and new canonic sensitivity schemes can also be obtained. The new *generic optimal controller scheme* seems to be the best among the investigated methods. It behaves also well because the identification and control errors are the same and they can vanish only at the same time. This scheme is easy to be implemented because general identification techniques can be applied between a calculated auxiliary input and the measured closed-loop output variables.

An applicable strategy for iterative control refinement based on the *generic scheme* was presented and illustrated by simulation examples. The adaptive version of the control refinement strategy was also shown and demonstrated by an example.

Acknowledgements

This work was supported in part by US ARO Contract #N68171-94-0-9064 and by the Hungarian NSF (OTKA)

References

[1] ANDERSON, B. D. O. – KOSUT, R. L. (1991): Adaptive Robust Control: On-line Learning, *30th IEEE Conf. Decision and Control,* Brighton, England.

[2] ASTRÖM, K. J. (1993): Matching Criteria for Control and Identification, *European Control Conference,* Groningen, NL, pp. 248–251.

[3] ASTRÖM, K. J. – WITTENMARK B. (1984): Computer Controlled Systems – Theory and Design. Prentice-Hall, Englewood Cliffs, N. J.

[4] BITMEAD, R. (1993): Iterative Control Design Approaches, *12th IFAC Congress,* Sydney, Australia, Vol. 9, pp. 381–384.

[5] BOKOR, J. – KEVICZKY, L. (1984): Structure and Parameter Estimation of MIMO Systems Using Elementary Subsystem Representation, *Int. J. of Control,* Vol. 39, 5, pp. 965–985.

[6] GEVERS, M. (1991): Connecting Identification and Robust Control: a New Challenge, *9th IFAC/IFORS Symposium on Identification and System Parameter Estimation,* Budapest, Hungary, pp. 1–10.

[7] HOROWITZ, I. M. (1963): Synthesis of Feedback Systems, Academic Press, New York.

[8] KEVICZKY, L. (1979): On the Transfer Functions of Sampled Continuous Systems, *Technical Report,* University of Minnesota, Department of Electrical Engineering, Minneapolis (USA).

[9] KEVICZKY, L. – BÁNYÁSZ, Cs. (1994): A New Structure to Design Optimal Control Systems. *IFAC Workshop on New Trends in Design of Control Systems.* pp. 102–105. Smolenice.

[10] LANDAU, I. D. (1990): System Identification and Control Design. Prentice- Hall, N. J.

[11] LEE, W. S. – ANDERSON, B. D. O. (1993): A New Approach to Adaptive Robust Control, *Int. J. of Adaptive Control and Signal Processing,* Vol. 7. pp. 183–211.

[12] LEE, W. S. – ANDERSON, B. D. O. – KOSUT, R. L. – MAREELS, I. M. Y. (1992): On Adaptive Robust Control and Control-relevant System Identification, *American Control Conference,* Chicago, USA, pp. 2834–2841.

[13] MORARI, M – ZAFIRIOU, E. (1989): Robust Process Control. Prentice-Hall international, inc. London.

[14] VAN DEN HOF, P. M. J. – SCHRAMA, R. J. P. – BOSGRA O. H. – DE CALLAFON, R. A. (1993): Identification of Normalized Coprime Plant Factors for Iterative Model and Controller Enhancement, *32nd IEEE Conf. Decision and Control,* San Antonio, TX, USA.

INDEX

INFORMATION FOR AUTHORS

Submitting a Manuscript for Publication. Submission of a paper to this journal implies that it represents original work previously not published elsewhere, and that it is not being considered elsewhere for publication. If accepted for publication, the copyright is passed to the publisher: the paper must not be published elsewhere in the same form, in any language, without written consent of the executive editor. The first author will receive 50 free reprints.

Manuscripts should be submitted in English in two copies to the editors' office (see inner front cover). Good office duplicated copies of the text are acceptable.

Periodica Polytechnica is typeset using the TEX program with the AMSTEX macro package. Therefore, authors are encouraged to submit their contribution after acceptance in this form too, or at least the text of the article in a simple ASCII file (e. g. via email or on a floppy diskette, readable on an IBM compatible PC). By this solution most of the typesetting errors can be avoided, and publishing time can be reduced. An AMSTEX style file with a sample TEX source file is also available upon request.

Compilation and Typing of Manuscripts. Contributions should be typed or printed in double spacing (24 pt spacing when using text processors), on A4 paper. One page may contain not more than 10 corrections (paints do not count).

The maximum length of the manuscript is 30 standard pages (25 lines, with 50 characters in a line), including illustrations and tables. When more characters are typed on a page, the allowed page number is reduced accordingly.

When using a text processor, please use a (preferably English) spelling checker before the final printing.

Use one side of the sheet only. Paragraphs are to be indented by 5 spaces and not to be preceded by a blank line.

A correctly compiled manuscript should contain the following items:

(1) Title page giving: the suggested running header (max. 50 characters) for the odd typeset pages; the title (short, with subheading if necessary); the name(s) of the author(s); affiliation (institution, firm etc.) of the author(s), in English, with mailing address, telefax and phone numbers and email address; grants, scholarships etc. during the research (in a footnote); an informative abstract of not more than 200 words with 3-5 keywords;
(2) Textual part of the contribution with a list of references;
(3) A separate sheet listing all figure captions and table headers;
(4) Illustrations and tables (please put your name on each sheet), at least one set of illustrations in very good quality for reproduction.

Abstract. A summary or abstract of about 100-200 words should be provided on the title page. This should be readable without reference to the article or to the list of references, and should indicate the scope of the contribution, including the main conclusions and essential original content.

Keywords are to be given for the purpose of data bases and abstracting; avoid too general keywords which provide no help in literature searching.

General rules for the text. Chapters are to be numbered with Arabic numbers in decimal hierarchy.

Wherever possible, mathematical equations should be typewritten, with subscripts and superscripts clearly shown. Metric (SI) units are to be used, other units may be given in parentheses. Equations must be numbered on the right side, in parentheses. Handwritten or rare mathematical, Greek and other symbols should be identified or even explained if necessary in the margin. Letters denoting quantities are to be distinguished by different setting both in the formulae and in the text. Remember the rule that scalar quantities are to be denoted by italics (underline by hand in your manuscript), vectors by lower case bold type letters (underline doubled), and matrices by bold capitals (underline doubled). Dimensions (like cm, Ohm, V etc.) and standard function names (sin, ln, P etc.) are to be typeset in Roman typefaces (not in italics). A few important words may be distinguished by italic setting (underline).

Illustrations and Tables. Graphs, charts and other line illustrations should be drawn neatly in Indian ink, or printed by a laser printer. Computer printouts can only be used if they are of excellent quality. Figures should be submitted in an adequate size for camera-ready pages in size 1.2:1. Suggested line thicknesses: 0.18–0.35–0.4 mm or 0.5–0.7–1.14 pt. Letter sizes: 0.4 mm (10 pt). All figures should be numbered with consecutive Arabic numbers, have descriptive captions, and be referred to in the text. Captions should be self-explanatory, not merely labels. Figures must not contain lengthy texts; use captions instead.

Number tables consecutively with Arabic numbers and give each a descriptive caption at the top. If possible avoid vertical rules in tables. Tables should be preferably submitted in camera-ready form.

The earliest acceptable position of figures and tables is to be indicated in the left margin.

References. In the text references should be made by the author's surname and year of publication in parentheses, e.g. (Earl et al, 1988) or ...was given by Kiss and Small (1986a). Where more than one publication by an author in one year is referred to, the year should be followed by a suffix letter (1986a, 1986b etc.), the same suffix being given in the reference list. For the style of the reference list, which is to be given in alphabetical order, see the examples below for journal articles, conference papers and books.

Earl, J., Kis, I. and Tōrōk, I. (1988): Partial Discharge Measurement in Cables. *Periodica Polytechnica Ser. Electrical Engineering*, Vol. 32, No. 4, pp. 133–138.

Kiss, S. and Small, A. B. (1986a): Roundoff Errors in FFT. *Proc. 5th IEEE Symposium on Signal Processing*, Boston (MA), May 3–5, 1986. New York, NY, IEEE Press, CH0092-2875/86, pp. 3.5–3.9.

Kiss, S. and Small, A. B. (1986b): Ellenállások (Resistances). Budapest, Tankönyvkiadó. pp. 533–535. (in Hungarian)

More detailed guidelines for authors, with hints for the preparation of figures and with a sample page, are available from the editors' office (see inner front cover).

CONTENTS